



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

DISTRIBUCIONES MATRIZ EXPONENCIALES

TESIS

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIO

PRESENTA:

MARIO EDGAR PIMENTEL ROSAS

DIRECTOR DE TESIS:

DR. MOGENS BLADT PETERSEN



2008



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

“Los años de búsqueda en la oscuridad de una verdad que uno siente pero no puede expresar, el deseo intenso y la alternancia de confianza y desazón, hasta que uno encuentra el camino a la claridad y comprensión, solo son familiares a aquel que los ha experimentado”

Albert Einstein

A mi abuelita

Por tu amor infinito, toda la sabiduría y enseñanzas que nos dejaste.
Gracias en donde quiera que estés.

A mis padres

Por su amor, ternura y comprensión que a lo largo de esta vida me han brindado.
Gracias por guiarme en esta vida y hacerme un hombre de bien.

A Paola

Por ser la persona más importante en esta vida, por todo el amor que me has brindado.
Por todo el apoyo incondicional en este sueño que has emprendido a mi lado.
Gracias, creo que faltan las palabras para expresar lo mucho que te amo.

A mis hermanas

Por ser el ejemplo a seguir y por todas las enseñanzas que me han brindado.

A mis tíos Alma y Gabriel

Por todo el apoyo incondicional que he recibido de su parte. Los quiero mucho.

A mis amigos

Gracias a todos y cada uno de ustedes por todo el cariño, confianza y amistad.
Por formar parte de mi vida y estar siempre a mi lado.

A mis profesores

Gracias a ustedes por enseñarme el maravilloso mundo de las matemáticas.

1. Datos del alumno
Pimentel
Rosas
Mario Edgar
56 72 71 19
Universidad Nacional Autónoma de México
Facultad de Ciencias
Actuario
300158193
2. Datos del tutor
Dr
Mogens
Bladt
Petersen
3. Datos del sinodal 1
Dr
Alberto
Contreras
Cristán
4. Datos del sinodal 2
Eduardo
Gutiérrez
Peña
5. Datos del sinodal 3
Ramsés
Mena
Chávez
6. Datos del sinodal 4
Pablo
Padilla
Longoria
7. Datos del trabajo escrito
Distribuciones matriz exponenciales
106 p.
2008

Contenido

Introducción	1
Capítulo 1: Conceptos Preliminares	
1.1 Introducción	3
1.2 El proceso de riesgo	3
1.3 Métodos	7
1.3.1 Soluciones Exactas	7
1.3.2 Métodos Numéricos	8
1.3.3 Inversión de la Transformada de Laplace	8
1.3.4 Métodos Analíticos Matriciales	8
1.3.5 Ecuaciones Diferenciales e Integrales	9
1.3.6 Aproximaciones	9
1.3.6.1 Aproximación de Cramér-Lundberg	9
1.3.6.2 Aproximaciones de Difusión	9
1.3.6.3 Aproximaciones de Reclamaciones Grandes	9
1.3.7 Límites y Desigualdades	9
1.3.8 Métodos Estadísticos	10
1.3.9 Simulación	10
1.4 La Transformada de Laplace	10
1.5 Ecuaciones Diferenciales de Kolmogorov	12
1.6 Álgebra Matricial	15
Capítulo 2: Procesos de Markov	
2.1 Introducción	20
2.2 Construcción Minimal	22
2.3 La Matriz de Intensidad	24
2.3.1 Definición y Unicidad	24
2.3.2 Ecuaciones Diferenciales de Kolmogorov	25

2.4 Estacionalidad y resultados límite	27
2.4.1 Clasificación de Estados	27
2.4.2 Medidas Estacionarias	27
Capítulo 3: Distribuciones Tipo Fase	
3.1 Introducción	30
3.2 Definición y Propiedades Básicas de las Distribuciones Tipo Fase	30
3.3 Representaciones Irreducibles	35
3.4 Propiedades de Cerradura de las Distribuciones Tipo Fase	37
3.5 Reclamaciones Tipo Fase	43
3.6 Ejemplos de Distribuciones Tipo Fase	46
3.7 Distribuciones Tipo Fase Multivariadas	48
Capítulo 4: Distribuciones Matriz Exponenciales	
4.1 Introducción	50
4.2 Definición y Propiedades Básicas de las Distribuciones Matriz Exponenciales	50
4.3 Funciones de Distribución	66
4.4 Teoría de Renovación	70
4.4.1 Diagonalización	74
Capítulo 5: Interpretación Física de las Distribuciones Matriz Exponenciales	
5.1 Introducción	78
5.2 Cálculo con Flujos: Teoría de Renovación	80
5.3 Modelo de Teoría del Riesgo	83
5.4 Procesos de Arribo Racionales	86
Capítulo 6: Distribuciones Matriz Exponenciales Multivariadas	
6.1 Fracciones Continuas	89
6.2 Distribuciones Matriz Exponenciales Multivariadas	94
6.3 Distribuciones Matriz Exponenciales Multivariadas y Teoría de Renovación	101
Conclusiones	103
Bibliografía	105

Introducción

Esta tesis muestra los resultados acerca de la clase de las distribuciones matriz exponenciales, destacando principalmente su definición equivalente como la clase de las distribuciones que tienen transformada de Laplace racional, para poder describir los modelos estocásticos de flujos como distribuciones matriz exponenciales y determinar la probabilidad de ruina en un modelo de riesgo con distribuciones matriz exponenciales.

Las distribuciones matriz exponenciales son una clase más amplia que las distribuciones tipo fase; éstas tienen funciones de distribución de la misma forma que las distribuciones tipo fase pero sus representaciones no necesariamente tienen una interpretación probabilística simple. Por este motivo es necesario describir las distribuciones tipo fase, que basan su análisis en modelar intervalos aleatorios de tiempo como un número de segmentos distribuidos exponencialmente, y aprovechar la estructura Markoviana para simplificar el análisis, así como describir los procesos de Markov.

Una variable aleatoria que es definida como el tiempo de paro de la absorción de una cadena de Markov absorbente a tiempo continuo se dice que tiene una distribución tipo fase. Las distribuciones y funciones de densidad de una distribución tipo fase pueden ser expresadas en términos de un vector de distribución inicial α de tamaño $1 \times m$ y una matriz generadora infinitesimal \mathbf{T} de tamaño $m \times m$ de la cadena de Markov. Los parámetros (α, \mathbf{T}) son conocidos como una representación de orden m de la distribución tipo fase.

Las distribuciones tipo fase, así como las distribuciones matriz exponenciales, pertenecen a la rama de la probabilidad computacional conocida como métodos analíticos matriciales. La probabilidad computacional fue descrita por Neuts (1981) como el estudio de los modelos estocásticos con un interés particular por la viabilidad algorítmica sobre un amplio rango de parámetros y valores.

Los métodos analíticos matriciales nos llevan al análisis de modelos estocásticos, particularmente sistemas de colas, usando métodos matriciales para desarrollar soluciones algorítmicamente tratables. El crecimiento de esta área se debe en gran parte a la capacidad de las computadoras para realizar cálculos numéricos más sofisticados.

Aunque las distribuciones matriz exponenciales no pertenecen en un estricto sentido a los métodos analíticos matriciales, los modelos estocásticos que utilizan distribuciones matriz exponenciales en lugar de distribuciones tipo fase tienen mayor flexibilidad y generalidad pero no siempre tienen interpretaciones probabilísticas simples.

Previo a las computadoras, algunos problemas en modelos estocásticos estaban relacionados con la transformada de Laplace y los métodos de análisis complejo para su solución.

Gracias a las computadoras, el campo de la probabilidad computacional así como también los métodos analíticos matriciales han redefinido el significado de una solución a un problema en los modelos estocásticos: un algoritmo implementable en el sistema que se está analizando. Por lo tanto, el número de sistemas que ahora pueden ser modelados estocásticamente se ha incrementado significativamente.

Durante las últimas dos décadas ha habido un incremento en la teoría y la aplicación de los métodos analíticos matriciales. Debido a esto, la complejidad de los modelos estocásticos que pueden ser analizados también ha crecido considerablemente.

Las áreas de aplicación han incluido planeación, teoría del riesgo, mantenimiento de máquinas, análisis de supervivencia y cinética de los medicamentos. La mayor actividad de investigación, dada la explosión en el tráfico de datos durante los últimos años, ha sido indudablemente en

análisis del funcionamiento de sistemas de telecomunicaciones. La literatura acerca de telecomunicaciones y la ingeniería electrónica ha sido inundada con aplicaciones de métodos analíticos matriciales.

Debido a esto, este trabajo se desarrolla en seis capítulos de la siguiente manera:

En el capítulo 1 se definen y se discuten algunos conceptos preliminares como los procesos de riesgo, los principales métodos que se utilizan y el proceso de superávit de reclamaciones; también se define formalmente la transformada de Laplace y se describen sus propiedades principales; se describen las ecuaciones diferenciales de Kolmogorov; por último se describen algunos resultados de álgebra matricial.

En el capítulo 2 se definen los procesos de Markov; se describen sus principales propiedades tales como la estructura básica, la construcción minimal y la matriz de intensidad entre otras.

En el capítulo 3 se definen las distribuciones tipo fase como la distribución del tiempo de paro hasta la absorción para un proceso de Markov finito de saltos; se describen sus propiedades, representaciones y caracterizaciones; así como también las distribuciones tipo fase multivariadas.

En el capítulo 4 se definen las distribuciones matriz exponenciales como las distribuciones con transformada de Laplace racional; se describen sus propiedades, representaciones y caracterizaciones; también se describen claramente los modelos estocásticos de flujos como distribuciones matriz exponenciales; se describen los procesos de renovación donde la distribución entre los arribos es matriz exponencial y finalmente se describe un modelo de riesgo donde la distribución de las reclamaciones es matriz exponencial.

En el capítulo 5 se describen claramente los modelos estocásticos de flujos como distribuciones matriz exponenciales; se describen los procesos de renovación donde la distribución entre los arribos es matriz exponencial y finalmente se describe un modelo de riesgo donde la distribución de las reclamaciones es matriz exponencial.

En el capítulo 6 se describe la clase de las distribuciones matriz exponenciales multivariadas como una extensión natural de la definición de las distribuciones matriz exponenciales univariadas y son definidas como las distribuciones en \mathbb{R}_+^n que tienen transformada de Laplace racional multidimensional; además se describen las funciones racionales en términos de fracciones continuas, así como también las matrices de Hankel.

Capítulo 1: Conceptos Preliminares

1.1 Introducción

En este capítulo se definen y se discuten algunos conceptos preliminares relacionados con la teoría que se presenta a lo largo del presente trabajo, como los procesos de riesgo y el proceso de superávit de reclamaciones, se define formalmente la transformada de Laplace (Assmusen, 2000) y sus propiedades principales, se describen las ecuaciones diferenciales de Kolmogorov y por último algunos resultados acerca del álgebra matricial (Rolski et al. 1999).

1.2 El Proceso de Riesgo

El estudio moderno de las probabilidades de ruina, también conocido como teoría del riesgo, se inició en Suecia en la primera mitad del siglo XX.

La escuela sueca no solo fue pionera en teoría del riesgo, sino también en probabilidad aplicada. En particular muchos resultados y métodos sobre caminatas aleatorias se originan en esta escuela y también en las áreas relacionadas, como la teoría de colas.

El término teoría de riesgo es interpretado en un sentido amplio como el estudio de las probabilidades de ruina de una compañía aseguradora.

Un proceso de riesgo $\{R_t\}_{t \geq 0}$, es un modelo para la evolución de la reserva de una compañía aseguradora en el tiempo. La probabilidad de ruina $\psi(u)$ es la probabilidad de que la reserva de la compañía aseguradora con una reserva inicial u al tiempo $t \geq 0$ sea menor que 0. Esto es

$$\psi(u) = \mathbb{P}\left(\inf_{t \geq 0} R_t < 0 \mid R_0 = u\right) \quad (1.2.1)$$

La probabilidad de ruina antes del tiempo T es

$$\psi(u, T) = \mathbb{P}\left(\inf_{0 \leq t \leq T} R_t < 0\right). \quad (1.2.2)$$

Nos referimos a (1.2.1) y (1.2.2) como las probabilidades de ruina de horizonte infinito y horizonte finito respectivamente.

Frecuentemente es más conveniente trabajar con el proceso de superávit $\{S_t\}_{t \geq 0}$ definido por

$$S_t = u - R_t.$$

Definimos el tiempo de ruina como

$$\tau(u) = \inf\{t \geq 0 : R_t < 0\} = \inf\{t \geq 0 : S_t > u\},$$

el máximo con horizonte infinito como

$$M = \sup_{0 \leq t < \infty} S_t$$

y el máximo con horizonte finito como

$$M_T = \sup_{0 \leq t < T} S_t.$$

Las probabilidades de ruina pueden ser escritas alternativamente como:

$$\psi(u) = \mathbb{P}(\tau(u) < \infty) = \mathbb{P}(M > u),$$

$$\psi(u, T) = \mathbb{P}(\tau(u) < T) = \mathbb{P}(M_T > u).$$

Existe sólo un número finito de reclamaciones en intervalos de tiempo finitos. Esto es, el número N_t de arribos en $[0, t]$ es finito. Denotamos a los tiempos entre los arribos de las reclamaciones por T_2, T_3, \dots ; es decir, T_1 es el tiempo de la primera reclamación. Así, el tiempo de arribo de la n -ésima reclamación es

$$\sigma_n = T_1 + \dots + T_n$$

y

$$N_t = \min\{n \geq 0 : \sigma_{n+1} > t\} = \max\{n \geq 0 : \sigma_n \leq t\}.$$

El tamaño de la n -ésima reclamación es denotada por U_n .

Las primas fluyen a una tasa p (prima de tarifa), por unidad de tiempo; por lo tanto,

$$R_t = u + pt - \sum_{k=1}^{N_t} U_k$$

y

$$S_t = \sum_{k=1}^{N_t} U_k - pt.$$

Los modelos que consideramos tienen generalmente la propiedad de que existe una constante ρ tal que

$$\frac{1}{t} \sum_{k=1}^{N_t} U_k \xrightarrow{c.s.} \rho, \quad t \rightarrow \infty, \quad (1.2.3)$$

donde ρ se interpreta como el monto promedio de las reclamaciones por unidad de tiempo. Una cantidad básica es la carga de seguridad η definida como el monto relativo por el cual la prima de tarifa excede a ρ

$$\eta = \frac{p - \rho}{\rho}.$$

La compañía aseguradora debe tratar de garantizar que $\eta > 0$.

Proposición 1.2.1: Supongamos que (1.2.3) ocurre. Si $\eta < 0$, entonces $M = \infty$ casi seguramente y por lo tanto $\psi(u) = 1$ para toda u . Si $\eta > 0$ entonces $M < \infty$ casi seguramente y por lo tanto $\psi(u) < 1$ para toda u suficientemente grande.

Demostración: Se sigue por (1.2.3) que

$$\frac{S_t}{t} = \frac{\sum_{k=1}^{N_t} U_k - pt}{t} \xrightarrow{c.s.} \rho - p, \quad t \rightarrow \infty.$$

Si $\eta < 0$, entonces este límite es mayor que cero lo cual implica $S_t \xrightarrow{c.s.} \infty$ y por lo tanto $M = \infty$ casi seguramente. Si $\eta > 0$, entonces $\lim \frac{S_t}{t} < 0$, $S_t \xrightarrow{c.s.} -\infty$, y $M < \infty$ casi seguramente. ■

Las clases más populares de distribuciones B que son utilizadas para modelar las reclamaciones U_1, U_2, \dots pueden clasificarse en dos grupos: las distribuciones de colas ligeras donde la función generadora de momentos es finita; es decir, $\hat{B}[s] < \infty$ para alguna $s > 0$; y las distribuciones de colas pesadas donde la función generadora de momentos es infinita; es decir, $\hat{B}[s] = \infty$ para toda $s > 0$.

Las distribuciones con las que trabajaremos a lo largo del presente trabajo son distribuciones de colas ligeras.

Una variable aleatoria exponencial es aquella cuya función de densidad está dada, para alguna $\lambda > 0$, por

$$f(x) = \lambda e^{-\lambda x} \quad \text{si } x \geq 0.$$

La función de distribución está dada por

$$F(x) = 1 - e^{-\lambda x}.$$

Las variables aleatorias exponenciales son las más simples para hacer frente a la teoría de riesgo.

La característica fundamental de las variables aleatorias exponenciales es la propiedad de la pérdida de memoria.

$$\mathbb{P}(X > s+t | X > t) = \mathbb{P}(X > s) \quad \text{para toda } s, t \geq 0. \quad (1.2.4)$$

La condición (1.2.4) es equivalente a

$$\frac{\mathbb{P}(X > s+t, X > t)}{\mathbb{P}(X > t)} = \mathbb{P}(X > s)$$

ó

$$\mathbb{P}(X > s+t) = \mathbb{P}(X > s)\mathbb{P}(X > t).$$

■

Una variable aleatoria gamma con parámetros (α, λ) , $\alpha > 0$, $\lambda > 0$ cuya función de densidad está dada por

$$f(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \text{ si } x \geq 0,$$

donde $\Gamma(\alpha)$ es la función gamma, definida como

$$\Gamma(\alpha) = \int_0^{\infty} e^{-x} x^{\alpha-1} dx.$$

La función generadora de momentos está dada por

$$\hat{B}[s] = \left(\frac{\lambda}{\lambda - s} \right)^{\alpha}.$$

La esperanza de una variable aleatoria gamma está dada por

$$\mathbb{E}(X) = \frac{\alpha}{\lambda}$$

y la varianza

$$\text{Var}(X) = \frac{\alpha}{\lambda^2}.$$

La forma exacta de la cola $\bar{B}(x)$ está dada por la función gamma incompleta $\Gamma(x, \alpha)$,

$$\bar{B}(x) = \frac{\Gamma(\lambda x, \alpha)}{\Gamma(\alpha)} \text{ donde } \Gamma(x, \alpha) = \int_x^{\infty} t^{\alpha-1} e^{-t} dt.$$

Si α es entero y X es una variable aleatoria gamma (α, λ) , entonces $X \stackrel{D}{=} X_1 + \dots + X_{\alpha}$ donde X_1, X_2, \dots son variables aleatorias exponenciales independientes idénticamente distribuidas con parámetro λ . Este caso especial de una variable aleatoria gamma también se conoce como una variable aleatoria Erlang con parámetro α .

Una distribución hiperexponencial generalizada tiene función de distribución, definida para $u \geq 0$, de la forma

$$F(u) = \sum_{i=1}^n a_i (1 - e^{-\lambda_i u}),$$

donde a_1, a_2, \dots, a_n son todos reales con $\sum_{i=1}^n a_i = 1$, y $\lambda_1 > \lambda_2 > \dots > \lambda_n > 0$.

Una distribución tipo fase es la distribución del tiempo de absorción en un proceso de Markov con espacio de estados finito, de los cuales un estado es absorbente y los demás estados son transitorios. (Ver Capítulo 3)

Los casos más importantes son las distribuciones exponenciales, Erlang e hipereponenciales. Esta clase de distribuciones juega un papel importante dentro de las formas exactas computacionalmente tratables de la probabilidad de ruina $\psi(u)$ que pueden ser obtenidas.

Los parámetros de una distribución tipo fase son el conjunto E de estados transitorios, la matriz de intensidad T del proceso de Markov y el vector renglón $\alpha = (\alpha_i)_{i \in E}$ de probabilidades iniciales.

Las distribuciones B tienen transformada de Laplace racional si $\hat{B}[s] = \frac{p(s)}{q(s)}$ con $p(s)$ y $q(s)$ polinomios de grado finito. Las caracterizaciones equivalentes consisten en que las densidades $b(x)$ tienen las siguientes formas:

$$b(x) = \sum_{j=1}^q c_j x^j e^{\eta_j x} \quad (1.2.5)$$

$$b(x) = \sum_{j=1}^{q_1} c_j x^j e^{\eta_j x} + \sum_{j=1}^{q_2} d_j x^j \cos(a_j x) e^{\delta_j x} + \sum_{j=1}^{q_3} e_j x^j \sin(b_j x) e^{\zeta_j x} \quad (1.2.6)$$

donde los parámetros en (1.2.5) pueden ser valores complejos mientras que los parámetros en (1.2.6) son valores reales.

Esta clase de distribuciones es popular en teoría del riesgo y teoría de colas, pero la tendencia actual en probabilidad aplicada es restringir la atención a las distribuciones tipo fase (Ver Capítulo 3) y a las distribuciones matriz exponenciales (Ver Capítulo 4).

1.3 Métodos

La teoría del riesgo puede ser vista como una de las diversas áreas de la probabilidad aplicada. Otras áreas son: procesos de ramificación, modelos genéticos, teoría de colas, procesos de almacenaje, confiabilidad, sistemas de interacción de partículas, ecuaciones diferenciales estocásticas, series de tiempo, procesos Gaussianos, teoría de valores extremos, geometría estocástica, procesos puntuales, etc.

Las áreas más relacionadas con la teoría del riesgo son la teoría de colas y los procesos de almacenaje.

1.3.1 Soluciones Exactas

Lo ideal es poder encontrar soluciones de forma cerrada para las probabilidades de ruina $\psi(u)$, $\psi(u, T)$. Los casos para la probabilidad de ruina de horizonte infinito $\psi(u)$ donde esto es posible son básicamente los siguientes:

El modelo Poisson compuesto con una tasa constante $p=1$ y con una distribución exponencial B para el tamaño de las reclamaciones.

El modelo Poisson compuesto con una tasa constante $p=1$ y con una distribución B tipo fase para el tamaño de las reclamaciones. En este caso $\psi(u)$ está dada en términos de una función matriz exponencial, la cual puede ser expandida en una suma de términos exponenciales por diagonalización.

El modelo Poisson compuesto con una distribución para el tamaño de las reclamaciones degenerada en un punto.

El modelo Poisson compuesto con una tasa $p(x)$ dependiendo de la reserva y con una distribución exponencial B para el tamaño de las reclamaciones.

Un proceso de Levý α estable con flujo donde $\psi(u)$ está dada como una serie infinita involucrando a la función de Mittag Leffer.

También los modelos Brownianos de caminatas aleatorias con ciertos saltos de libertad nos conducen a soluciones explícitas. Un hecho notable es la forma explícita de la probabilidad de ruina cuando $\{R_t\}$ es un proceso de difusión con flujo infinitesimal $\mu(x)$ y varianza $\sigma^2(x)$

$$\psi(u) = \frac{\int_u^\infty \exp\left\{-\int_0^x 2\frac{\mu(y)}{\sigma^2(y)} dy\right\}}{\int_0^\infty \exp\left\{-\int_0^x 2\frac{\mu(y)}{\sigma^2(y)} dy\right\}} = 1 - \frac{S(u)}{S(\infty)},$$

donde

$$S(u) = \int_0^u \exp\left\{-\int_0^x 2\frac{\mu(y)}{\sigma^2(y)} dy\right\},$$

es la escala natural.

Para la probabilidad de ruina de horizonte finito $\psi(u, T)$, el ejemplo de una expresión explícita es el modelo Poisson compuesto con una tasa constante $p=1$ y con una distribución exponencial para el tamaño de las reclamaciones.

1.3.2 Métodos Numéricos

Además de una solución de forma cerrada, la mejor alternativa después de esta es un procedimiento numérico que nos permita calcular los valores exactos de las probabilidades de ruina.

1.3.3 Inversión de la Transformada de Laplace

A menudo es más fácil encontrar las transformadas de Laplace

$$\hat{\psi}(-s) = \int_0^\infty e^{-su} \psi(u) du, \quad \hat{\psi}(-s, -\omega) = \int_0^\infty \int_0^\infty e^{-su - \omega T} \psi(u, T) dudT$$

en forma cerrada que las probabilidades de ruina por sí mismas. Considerando que esto puede llevarse a cabo, $\psi(u)$, $\psi(u, T)$ pueden ser calculadas entonces numéricamente por algún método de inversión de la transformada de Laplace.

1.3.4 Métodos Analíticos Matriciales

Este enfoque es relevante cuando la distribución del tamaño de la reclamación es tipo fase o matriz exponencial (distribuciones con transformada de Laplace racional), y en algunos casos la probabilidad de ruina $\psi(u)$ está dada por una función matriz exponencial $e^{\mathbf{U}u}$ que puede ser calculada por diagonalización, como la solución de ecuaciones lineales diferenciales o por alguna expansión en serie. En el proceso Poisson compuesto, \mathbf{U} es explícito en términos de los parámetros del modelo, mientras que para el modelo con arribos de renovación o los modelos derivados de procesos de Markov, \mathbf{U} se tiene que calcular numéricamente como la

solución iterativa de un problema de punto fijo o encontrando la forma diagonal en términos de las raíces complejas de ciertas ecuaciones trascendentales.

1.3.5 Ecuaciones Diferenciales e Integrales

La idea de este enfoque es expresar la probabilidad de ruina (1.2.1) o (1.2.2) como la solución de una ecuación diferencial o integral, y obtener una solución por algún método numérico. Un ejemplo donde podemos obtener una solución mediante una adecuada elección de variables es el caso donde las reclamaciones son de tipo fase.

1.3.6 Aproximaciones

1.3.6.1 Aproximación de Cramér-Lundberg

Este es uno de los resultados más importantes de la teoría del riesgo. Para el modelo Poisson compuesto con $p = 1$ y con una distribución para el tamaño de las reclamaciones B con función generadora de momentos $\hat{B}[s]$, señala que

$$\psi(u) \sim Ce^{-\gamma u}, \quad u \rightarrow \infty. \quad (1.3.6.1)$$

Donde $C = \frac{(1-\rho)}{(\beta - \hat{B}[\gamma])}$ y $\gamma > 0$ es la solución de la ecuación de Lundberg

$$(\beta - \hat{B}[\gamma]) - \gamma = 0.$$

Frecuentemente γ es llamado el coeficiente de ajuste. La aproximación de Cramér-Lundberg tiene generalizaciones para los modelos con arribos de renovación en un ambiente Markoviano. Sin embargo, en algunos casos la evaluación de C es más complicada. De hecho, cuando la distribución del tamaño de las reclamaciones es tipo fase, la solución exacta es tan sencilla de calcular como la aproximación de Cramér-Lundberg.

1.3.6.2 Aproximaciones de Difusión

Aquí la idea es aproximar el proceso de riesgo mediante un movimiento Browniano ajustando el primer y segundo momentos, y usar el hecho de que las primeras probabilidades de transición son calculadas más fácilmente por difusiones que por los procesos de riesgo. Las aproximaciones de difusión son sencillas de calcular pero no son muy precisas en su implementación. Sin embargo, las aproximaciones de difusión corregidas son por mucho lo mejor que se puede hacer en términos de probabilidades de ruina de horizonte finito $\psi(u, T)$.

1.3.6.3 Aproximaciones de reclamaciones grandes

Para que la aproximación de Cramér-Lundberg sea válida, la cola de la distribución del tamaño de las reclamaciones $\bar{B}(x)$ debe tener un decrecimiento exponencial. Las aproximaciones para $\psi(u)$ son tan precisas como para $\psi(u, T)$ para u suficientemente grandes.

1.3.7 Límites y desigualdades

El resultado principal en el área es la desigualdad de Lundberg

$$\psi(u) \leq e^{-\gamma u}.$$

Comparado con la aproximación de Cramér-Lundberg (1.3.6.1), esta tiene la ventaja de no involucrar aproximaciones y también como regla general, siendo algo más fácil generalizar más allá del ajuste del modelo Poisson compuesto.

1.3.8 Métodos estadísticos

Todos los métodos y resultados mencionados anteriormente asumen que los parámetros del modelo son completamente conocidos. En la práctica, los parámetros deben ser estimados de los datos observados en el proceso de riesgo en el intervalo de tiempo $[0, T]$. Este procedimiento es por sí mismo transparente. Sin embargo, la dificultad se presenta cuando tratamos de inferir acerca de las probabilidades de ruina. ¿Cómo podemos producir un intervalo de confianza? Y más importante aún, ¿Podremos confiar en los intervalos de confianza para valores grandes de u , los cuales son de interés? Según Assmussen (2000) esta extrapolación de los datos observados se debe a la extrema sensibilidad de las probabilidades de ruina de la cola de la distribución del tamaño de las reclamaciones. La sugerencia es observar que la media de la vida residual

$$\mathbb{E}[U - x | U > x] = \frac{1}{\bar{B}(x)} \int_0^{\infty} (y - x) B(dy)$$

normalmente tiene un límite finito en el caso de las colas ligeras y tiende a ∞ en el caso de las colas pesadas, y graficar la media empírica de la vida residual para observar algún comportamiento restrictivo que sea aparente en la cola de la distribución, es decir, graficar

$$\frac{1}{N - k} \sum_{i=k+1}^N (U_{(i)} - U_{(k)})$$

como función de $U_{(k)}$, donde $U_{(1)}, U_{(2)}, \dots, U_{(N)}$ son los estadísticos de orden basados en N variables aleatorias independientes idénticamente distribuidas que son las reclamaciones U_1, U_2, \dots, U_N para observar comportamientos aparentes en la cola.

1.3.9 Simulación

El desarrollo de las computadoras ha hecho a la simulación una herramienta popular en todas las ramas de la probabilidad aplicada y de la estadística, y desde luego este método es relevante en la teoría del riesgo. La simulación puede ser usada para obtener alguna idea acerca del proceso bajo estudio: simular algunas trayectorias muestrales, y observar cuáles de ellas exhiben el comportamiento esperado o algún comportamiento o resultado inesperado. Sin embargo, la situación más común es realizar una simulación Monte Carlo para estimar probabilidades, esperanzas o distribuciones que no están disponibles analíticamente.

El caso de las probabilidades de ruina de horizonte infinito presenta una dificultad, porque al parecer requiere una simulación infinita. Un truncamiento a un problema de horizonte finito ha sido usado, pero no es muy satisfactorio. Existen buenos métodos para un gran número de modelos y están basados en representar la probabilidad de ruina $\psi(u)$ como el valor esperado de una variable aleatoria la cual puede ser generada por una simulación. El problema es enteramente análogo a estimar características estables por simulación en teoría de colas, y de hecho métodos de otras áreas pueden ser usados frecuentemente en teoría de riesgo. Sin embargo, la dificultad principal es un evento con una probabilidad mínima de ocurrencia.

1.4 La Transformada de Laplace

La transformada de Laplace-Stieltjes de la función de distribución F en X , está definida de la siguiente manera:

$$\hat{l}(s) = \mathbb{E}(e^{-sX}) = \int_{-\infty}^{\infty} e^{-sx} dF(x).$$

Además de la transformada de Laplace-Stieltjes, consideraremos la transformada de Laplace

$$F(s) = \int_{-\infty}^{\infty} e^{-st} f(t) dt,$$

de una función $f : \mathbb{R} \rightarrow \mathbb{R}_+$.

Claramente si la función de distribución F es continua con densidad f , entonces su transformada de Laplace-Stieltjes es igual a la transformada de Laplace de f .

Una transformada integral de la función $g : \mathbb{R} \rightarrow \mathbb{R}_+$ es una función \tilde{g} de la forma

$$\tilde{g}(\theta) = \int_{-\infty}^{\infty} K(\theta, x)g(x)dx \text{ donde } \theta \in \mathbb{C},$$

donde una función dada g se transforma en otra función \tilde{g} por medio de una integral. Se dice que la función \tilde{g} es la transformada de g y la función K se llama kernel de la transformación.

La idea general es aplicar la relación $\tilde{g}(\theta) = \int_{-\infty}^{\infty} K(\theta, x)g(x)dx$ a fin de transformar un problema para g en un problema más sencillo para \tilde{g} , resolver este problema más simple y luego recuperar la función deseada g a partir de su transformada \tilde{g} .

La transformada de Laplace de g es definida por la función

$$\hat{g}(\theta) = \int_{-\infty}^{\infty} e^{-\theta x} g(x)dx \text{ donde } \theta \in \mathbb{C},$$

siempre que esta integral exista. En esta transformada se usa el kernel $K(\theta, x) = e^{-\theta x}$.

Frecuentemente, existen funciones g que están definidas en $[0, \infty)$ con transformada de Laplace

$$\hat{g}(\theta) = \int_0^{\infty} e^{-\theta x} g(x)dx.$$

Pensamos en \hat{g} como una función de una variable real θ .

Bajo ciertas condiciones tales como la existencia y la continuidad, la transformada de Laplace tiene las siguientes propiedades:

Inversión. La función g puede ser recuperada al conocer a \hat{g} mediante la fórmula de inversión

Convolución. Si $k(x) = \int_{-\infty}^{\infty} g(x-y)h(y)dy$ entonces $\hat{k}(\theta) = \hat{g}(\theta)\hat{h}(\theta)$

Diferenciación. Si $G : [0, \infty) \rightarrow \mathbb{R}$ y $g = \frac{dG}{dx}$ entonces $\theta\hat{G}(\theta) = \hat{g}(\theta) + G(0)$

1.5 Ecuaciones Diferenciales de Kolmogorov

En esta sección mostraremos la correspondencia uno a uno entre la función matriz de transición y sus matrices de intensidad.

Teorema 1.5.1: Para toda $i, j \in E$ y $h \geq 0$, las funciones de transición $p_{ij}(h)$ son diferenciables y satisfacen el siguiente sistema de ecuaciones diferenciales

$$p_{ij}^{(1)}(h) = \sum_{k \in E} q_{ik} p_{kj}(h). \quad (1.5.1)$$

Demostración: Sea $h' > 0$. Por la ecuación de Chapman-Kolmogorov tenemos que

$$\begin{aligned} p_{ij}(h+h') - p_{ij}(h) &= \sum_{k \in E} p_{ik}(h') p_{kj}(h) - p_{ij}(h) \\ p_{ij}(h+h') - p_{ij}(h) &= \sum_{k \neq i} p_{ik}(h') p_{kj}(h) + (p_{ii}(h') - 1) p_{ij}(h). \end{aligned}$$

Similarmemente

$$\begin{aligned} p_{ij}(h-h') - p_{ij}(h) &= p_{ij}(h-h') - \sum_{k \in E} p_{ik}(h') p_{kj}(h-h') \\ p_{ij}(h-h') - p_{ij}(h) &= - \sum_{k \neq i} p_{ik}(h') p_{kj}(h-h') - (p_{ii}(h') - 1) p_{ij}(h-h'). \end{aligned}$$

Dividiendo por h' , haciendo $h' \downarrow 0$ y usando la continuidad de $p_{ij}(h)$ obtenemos (1.5.1).

Por el Teorema 1.5.1

$$\lim_{h' \downarrow 0} \frac{1}{h'} \sum_{k \neq i} p_{ik}(h') p_{kj}(h) = \sum_{k \neq i} q_{ik} p_{kj}(h)$$

y

$$\lim_{h' \downarrow 0} \frac{1}{h'} \sum_{k \neq i} p_{ik}(h') p_{kj}(h-h') = \sum_{k \neq i} q_{ik} p_{kj}(h).$$

Las ecuaciones diferenciales en (1.5.1) son llamadas las ecuaciones diferenciales de Kolmogorov. En notación matricial (1.5.1) toma la forma

$$\mathbf{P}^{(1)}(h) = \mathbf{Q}\mathbf{P}(h) \quad (1.5.2)$$

para toda $h \geq 0$. De la misma forma como fue demostrada (1.5.1) se puede mostrar que

$$\mathbf{P}^{(1)}(h) = \mathbf{P}(h)\mathbf{Q} \quad (1.5.3)$$

para toda $h \geq 0$, la cual es la notación matricial de la ecuación diferencial de Kolmogorov. La condición inicial para ambas ecuaciones de Kolmogorov es $\mathbf{P}(0) = \mathbf{I}$

Suponemos que todas las matrices consideradas tienen dimensión $l \times l$ y los vectores tienen dimensión $1 \times l$. La convergencia de las sucesiones de matrices es de la siguiente manera: si $\{\mathbf{A}_n\}$ es una sucesión de matrices, $\mathbf{A}_n \rightarrow \mathbf{0}$ conforme $n \rightarrow \infty$ significa que $(\mathbf{A}_n)_{ij} \rightarrow 0$ para toda $i, j = 1, \dots, l$. Introducimos las siguientes normas: para un vector $\mathbf{x} = (x_1, \dots, x_l)$, definimos $\|\mathbf{x}\| = \sum_{i=1}^l |x_i|$ y, para una matriz $\mathbf{A} = (a_{ij})$ de $l \times l$ definimos $\|\mathbf{A}\| = \sum_{i,j=1,\dots,l} |a_{ij}|$. Notemos que para $h \in \mathbb{R}$ tenemos que $\|h\mathbf{A}\| = |h|\|\mathbf{A}\|$ y que $\|\mathbf{A}\| = 0$ si y solo si $\mathbf{A} = \mathbf{0}$. Es claro que $\mathbf{A}_n \rightarrow \mathbf{0}$ si y solo si $\|\mathbf{A}_n\| \rightarrow 0$. Además, para $a \geq 0$ y matrices arbitrarias \mathbf{A}, \mathbf{B}

$$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|, \quad \|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|, \quad \|a\mathbf{A}\| = a\|\mathbf{A}\| \quad (1.5.4)$$

Lema 1.5.1: La serie $\sum_{n=0}^{\infty} \frac{(h\mathbf{A})^n}{n!}$ converge uniformemente con respecto a $h \in [-h_0, h_0]$, para cada $h_0 > 0$.

Demostración: Sea $h \in [-h_0, h_0]$ y $m \in \mathbb{N}$. Por la desigualdad del triángulo, deducida de (1.5.4), tenemos que

$$\left\| \sum_{n=0}^{\infty} \frac{(h\mathbf{A})^n}{n!} - \sum_{n=0}^m \frac{(h\mathbf{A})^n}{n!} \right\| = \left\| \sum_{n=m+1}^{\infty} \frac{(h\mathbf{A})^n}{n!} \right\| \leq \sum_{n=m+1}^{\infty} \frac{h^n \|\mathbf{A}\|^n}{n!} \leq \sum_{n=m+1}^{\infty} \frac{h_0^n \|\mathbf{A}\|^n}{n!} < \varepsilon$$

para cada m suficientemente grande, uniformemente en $h \in [-h_0, h_0]$. La serie $\sum_{n=0}^{\infty} \frac{(h\mathbf{A})^n}{n!}$ es por lo tanto una función matricial bien definida y es continua con respecto a h en \mathbb{R} . Llamaremos a esta función la función matriz exponencial y la denotaremos por

$$\exp(h\mathbf{A}) = \mathbf{I} + h\mathbf{A} + \dots + \frac{(h\mathbf{A})^n}{n!} + \dots \quad (1.5.5)$$

Sea $\mathbf{A}(h)$ una función matricial tal que todas sus entradas son funciones diferenciables de h . Definimos la matriz derivada por

$$\mathbf{A}^{(1)}(h) = \lim_{h' \rightarrow 0} h'^{-1} (\mathbf{A}(h+h') - \mathbf{A}(h)).$$

Lema 1.5.2: La función matriz exponencial $\exp(h\mathbf{A})$ es diferenciable en los números reales y

$$\frac{d \exp(h\mathbf{A})}{dh} = \mathbf{A} \exp(h\mathbf{A}) = \exp(h\mathbf{A}) \mathbf{A}. \quad (1.5.6)$$

Demostración: Tenemos que

$$\frac{\exp((h+h')\mathbf{A}) - \exp(h\mathbf{A})}{h'} = \sum_{n=1}^{\infty} \frac{(h+h')^n - h^n}{h'} \frac{\mathbf{A}^n}{n!} = \sum_{n=1}^{\infty} nh^{n-1} \frac{\mathbf{A}^n}{n!} + h' \sum_{n=1}^{\infty} r_n(h, h') \frac{\mathbf{A}^n}{n!},$$

donde

$$0 \leq r_n(h, h') \leq n(n-1)(2h)^n, \quad (1.5.7)$$

para $|h'| \leq |h|$. El límite (1.5.7) es obtenido por una serie de Taylor de la función

$$g(x) = (h+x)^n - h^n. \text{ Así obtenemos } g(x) = xnh^{n-1} + \frac{x^2}{2}n(n-1)(h+\theta x)^{n-2}, \text{ donde } 0 \leq \theta \leq 1. \text{ Por lo tanto, haciendo } h' \rightarrow 0 \text{ obtenemos que}$$

$$\frac{d \exp(h\mathbf{A})}{dh} = \sum_{n=1}^{\infty} nh^{n-1} \frac{\mathbf{A}^n}{n!}.$$

Claramente

$$\sum_{n=1}^{\infty} nh^{n-1} \frac{\mathbf{A}^n}{n!} = \mathbf{A} \sum_{n=1}^{\infty} \frac{(h\mathbf{A})^n}{n!} = \sum_{n=1}^{\infty} \frac{(h\mathbf{A})^n}{n!} \mathbf{A}.$$

■

Dos matrices \mathbf{A} , \mathbf{A}' de tamaño $l \times l$ son conmutativas cuando $\mathbf{A}\mathbf{A}' = \mathbf{A}'\mathbf{A}$ y tenemos que para las matrices conmutativas

$$\exp(\mathbf{A} + \mathbf{A}') = \exp(\mathbf{A})\exp(\mathbf{A}'). \quad (1.5.8)$$

Lema 1.5.3: Sea \mathbf{Q} una matriz de tamaño $l \times l$ arbitraria tal que $q_{ij} \geq 0$ para $i \neq j$ y $\mathbf{Q}\mathbf{e}^t = \mathbf{0}$. Entonces la función matriz exponencial $(\exp(h\mathbf{Q}), h > 0)$ es una función matriz de transición que resuelve las ecuaciones diferenciales de Kolmogorov (1.5.2) y (1.5.3).

Demostración: Notemos que para $\mathbf{Q} = \mathbf{0}$ el enunciado es obvio. Supongamos ahora que $\mathbf{Q} \neq \mathbf{0}$. Primero verificaremos que $\exp(h\mathbf{Q})$ es una matriz estocástica para toda $h > 0$. Si $\mathbf{Q}\mathbf{e}^t = \mathbf{0}$, tenemos que $\mathbf{Q}^n \mathbf{e}^t = \mathbf{0}$ para toda $n = 1, 2, \dots$ y además $\exp(h\mathbf{Q})\mathbf{e}^t = \mathbf{e}^t$. Para demostrar que todas las entradas de $\exp(h\mathbf{Q})$ son no negativas, sea $a = \max\{-q_{ii} : i = 1, \dots, l\}$. Entonces, $\tilde{\mathbf{P}}$ definida por

$$\tilde{\mathbf{P}} = a^{-1}\mathbf{Q} + \mathbf{I}.$$

Es una matriz no negativa y entonces $\tilde{\mathbf{P}}\mathbf{e}^t = a^{-1}\mathbf{Q}\mathbf{e}^t + \mathbf{I}\mathbf{e}^t = \mathbf{e}^t$, la matriz $\tilde{\mathbf{P}}$ es estocástica. El que las entradas de $\exp(h\mathbf{Q})$ son no negativas se siguen de la representación

$$\exp(h\mathbf{Q}) = \exp(ah(\tilde{\mathbf{P}} - \mathbf{I})) = \sum_{n=0}^{\infty} \frac{(ah)^n (\tilde{\mathbf{P}})^n}{n!} e^{-ah}.$$

En esta igualdad la ecuación (1.5.8) ha sido usada junto con el hecho que $\exp(-ah\mathbf{I}) = e^{-ah}\mathbf{I}$. Además la ecuación (1.5.8) implica que $\exp(h\mathbf{Q})$ cumple con la ecuación de Chapman-Kolmogorov. Usando la ecuación (1.5.6) notamos que \mathbf{Q} es la matriz de intensidad de la función matriz de intensidad $\exp(h\mathbf{Q})$.

Teorema 1.5.2: La función matriz de transición $(\mathbf{P}(h), h \geq 0)$ puede ser representada por su matriz de intensidad \mathbf{Q} por medio de

$$\mathbf{P}(h) = \exp(h\mathbf{Q}).$$

Demostración: Por el Lema 1.5.3, $\mathbf{P}'(h) = \exp(h\mathbf{Q})$ es una solución a la ecuación hacia atrás de Kolmogorov y cumple las condiciones iniciales $\mathbf{P}'(0) = \mathbf{I}$. De la teoría de sistemas de ecuaciones diferenciales ordinarias lineales sabemos que esta solución es única. Así $\mathbf{P}'(h) = \mathbf{P}(h)$, para cada $h \geq 0$.

Si los valores propios $\theta_1, \dots, \theta_l$ de \mathbf{Q} son distintos, entonces la representación espectral de \mathbf{Q}^n puede ser usada para determinar la función matriz exponencial $\mathbf{P}(h) = \exp(h\mathbf{Q})$. En este caso tenemos que

$$\mathbf{P}(h) = \sum_{i=1}^l e^{h\theta_i} \phi_i^t \psi_i,$$

donde ϕ_i, ψ_i son los vectores propios por la derecha y por la izquierda correspondientes a θ_i .

Por lo tanto para cada par fijo $i, j \in E$ la función $p_{ij}(h)$ es idénticamente cero o siempre positiva en $(0, \infty)$. ■

Notemos que la suposición de un espacio de estados finito es fundamental para el resultado del teorema. Si el espacio de estados es infinito $E = \{1, 2, \dots\}$ puede haber muchas funciones matriz de transición correspondientes a una matriz de intensidad.

1.6 Álgebra Matricial

La función exponencial escalar e^{at} puede representarse por la serie de potencias

$$e^{at} = 1 + \sum_{n=1}^{\infty} \frac{a^n t^n}{n!},$$

que converge para toda t . Al reemplazar ahora el escalar a por la matriz constante \mathbf{A} de $n \times n$ y considerar la serie correspondiente, podemos obtener el desarrollo de $e^{\mathbf{A}t}$ por la serie de potencias.

$$e^{\mathbf{A}t} = \mathbf{I} + \sum_{n=1}^{\infty} \frac{\mathbf{A}^n t^n}{n!} = \mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2 t^2}{2!} + \dots + \frac{\mathbf{A}^n t^n}{n!}. \quad (1.6.1)$$

Cada término en la serie es una matriz $n \times n$. Es posible demostrar que cada componente de esta suma matricial converge para toda t cuando $n \rightarrow \infty$ porque $\mathbf{A}^n = O(n^k |\lambda|^n)$ para algún entero $k < p$, donde λ es el valor propio más grande en valor absoluto, es decir, $|\lambda| = \max\{|\mu| : \mu \in \text{sp}(\mathbf{A})\}$ y $\text{sp}(\mathbf{A})$ es el conjunto de todos los valores propios de \mathbf{A}

$$\text{sp}(e^{\mathbf{A}}) = \{e^\lambda : \lambda \in \text{sp}(\mathbf{A})\}.$$

Por lo tanto, la serie define con su suma una nueva matriz, que se denota por $e^{\mathbf{A}t}$; es decir

$$e^{\mathbf{A}t} = \mathbf{I} + \sum_{n=1}^{\infty} \frac{\mathbf{A}^n t^n}{n!}.$$

Al derivar término a término la serie se obtiene

$$\frac{d}{dt}(e^{\mathbf{A}t}) = \sum_{n=1}^{\infty} \frac{\mathbf{A}^n t^{n-1}}{(n-1)!} = \mathbf{A} \left(\mathbf{I} + \sum_{n=1}^{\infty} \frac{\mathbf{A}^n t^n}{n!} \right) = \mathbf{A} e^{\mathbf{A}t}.$$

Por lo tanto, $e^{\mathbf{A}t}$ satisface la ecuación diferencial

$$\frac{d}{dt}(e^{\mathbf{A}t}) = \mathbf{A} e^{\mathbf{A}t}.$$

Supongamos que las matrices son de dimensión $l \times l$ y los vectores son de dimensión $1 \times l$.

Denotaremos por $\theta_i = \theta_i(\mathbf{A})$, $i = 1, \dots, l$, a los valores propios de la matriz $\mathbf{A} = (a_{ij})_{i,j \in E}$; $E = \{1, \dots, l\}$.

Lema 1.6.1: Si $\mathbf{A}' = a\mathbf{A} + b\mathbf{I}$ para algunas constantes $a, b \in \mathbb{R}$, entonces

$$\theta_i(\mathbf{A}') = a\theta_i(\mathbf{A}) + b, \quad i = 1, \dots, l. \quad (1.6.2)$$

Lema 1.6.2: Sea \mathbf{A} no negativa. Entonces,

$$|\theta_i| \leq \min \left\{ \max_{i \in E} \sum_{j=1}^l a_{ij}, \max_{j \in E} \sum_{i=1}^l a_{ij} \right\}. \quad (1.6.3)$$

Demostración: Tenemos que $\theta_i \phi_i = \sum_{j=1}^l a_{ij} \phi_j$, $i = 1, \dots, l$, donde $\boldsymbol{\phi} = (\phi_1, \dots, \phi_l)$. Así

obtenemos que $|\theta_i| |\phi_i| \leq \sum_{j=1}^l a_{ij} \max_{k \in E} |\phi_k|$, $i = 1, \dots, l$, y en consecuencia

$$|\theta_i| \max_{i \in E} |\phi_i| \leq \max_{j=1}^l a_{ij} \max_{k \in E} |\phi_k|.$$

Así $|\theta_1| \leq \max_{i \in E} \sum_{j=1}^l a_{ij}$. La demostración de que $|\theta_1| \leq \max_{i \in E} \sum_{j=1}^l a_{ij}$ es similar porque \mathbf{A}^T tiene los mismos valores propios que \mathbf{A} . ■

La restricción dada en el Lema 1.6.2 puede ser utilizada para demostrar que para cierta clase de matrices, la parte real de todos los valores propios es no positiva. Es decir, que $\mathbf{B} = (b_{ij})$ es una matriz de subintensidad, si $b_{ij} \geq 0$ ($i \neq j$) y $\sum_{j=1}^l b_{ij} \leq 0$, donde para al menos una $i \in E$, $\sum_{j=1}^l b_{ij} < 0$.

Teorema 1.6.1: Si \mathbf{B} es una matriz de subintensidad, entonces $\theta_i(\mathbf{B}) = 0$ o $\Re(\theta_i(\mathbf{B})) < 0$, para cada $i = 1, \dots, l$.

Demostración: Sea \mathbf{B} una matriz de subintensidad y $c > \max_{i \in E} (-b_{ii})$. Entonces $\mathbf{B}' = \mathbf{B} + c\mathbf{I}$ es una matriz no negativa. Por el Lema 1.5.1, tenemos que $\theta_i(\mathbf{B}) = \theta_i(\mathbf{B}') - c$. Además por el Lema 1.6.2 tenemos que $|\theta_i(\mathbf{B}')| \leq c$. Entonces $|\theta_i(\mathbf{B}')| \geq |\theta_2(\mathbf{B}')| \geq \dots \geq |\theta_l(\mathbf{B}')|$. ■

Corolario 1.6.1: Una matriz de subintensidad \mathbf{B} es no singular si y solo si $\Re(\theta_i(\mathbf{B})) < 0$, para cada $i = 1, \dots, l$.

Demostración: Es suficiente notar que 0 no es un valor propio de \mathbf{B} si y solo si \mathbf{B} es no singular. ■

En el siguiente teorema se dará una fórmula para la representación de una función matriz exponencial $\exp(t\mathbf{A})$ de una matriz arbitraria \mathbf{A} .

Teorema 1.6.2: Sean $\theta_1, \dots, \theta_l$ los valores propios de \mathbf{A} . Entonces

$$\exp(\mathbf{A}t) = a_1(t)\mathbf{A}_1 + \dots + a_l(t)\mathbf{A}_l \quad (1.6.4)$$

donde $a_k(t)$, \mathbf{A}_k están dadas recursivamente por $a_1(t) = e^{\theta_1 t}$, $\mathbf{A}_1 = \mathbf{I}$ y $a_k(t) = \int_0^t e^{\theta_k(t-x)} a_{k-1}(x) dx$, $\mathbf{A}_k = (\mathbf{A}\theta_1 - \mathbf{I}) \dots (\mathbf{A}\theta_{k-1} - \mathbf{I})$ para $k = 2, \dots, l$

Demostración: Supongamos primero que $\mathbf{A} = \text{diag}(\boldsymbol{\theta})$ con $\boldsymbol{\theta} = (\theta_1, \dots, \theta_l)$ y todos sus valores propios $\theta_1, \dots, \theta_l$ son distintos. Entonces por la definición (1.6.1) de $\exp(\mathbf{A}t)$, tenemos que $\exp(\mathbf{A}t) = \text{diag}(e^{\theta_1 t}, \dots, e^{\theta_l t})$ y (1.6.4) es obvia para $l = 1$. Supongamos que (1.6.4) se cumple para alguna $l = n - 1$. Entonces el lado derecho (1.6.4) puede ser escrito como

$$\sum_{k=1}^n a_k(t)\mathbf{A}_k = \begin{pmatrix} \text{diag}(e^{\theta_1 t}, \dots, e^{\theta_{n-1} t}) & \mathbf{0} \\ \mathbf{0} & \sum_{k=1}^n a_k(t) \prod_{i=1}^{k-1} (\theta_n - \theta_i) \end{pmatrix}.$$

Tenemos que demostrar que, para toda $n = 1, 2, \dots$

$$\sum_{k=1}^n a_k(t) \prod_{i=1}^{k-1} (\theta_n - \theta_i) = e^{\theta_n t}. \quad (1.6.5)$$

Para $n = 1$, (1.6.5) es obvio. Además, para $k \geq 2$ tenemos que

$$a_k(t) = \int_0^t \int_0^{x_k} \dots \int_0^{x_3} e^{\theta_k(t-x_k)} \dots e^{\theta_2(x_3-x_2)} e^{\theta_1 x_2} dx_2 \dots dx_k$$

$$a_k(t) = e^{\theta_k t} \int_0^t \dots \int_0^t 1(0 < x_2 < \dots < x_k) \times e^{(\theta_{k-1}-\theta_k)x_k} \dots e^{(\theta_2-\theta_3)x_3} e^{(\theta_1-\theta_2)x_2} dx_2 \dots dx_k$$

$$a_k(t) = \int_0^t e^{\theta_1 x_2} e^{\theta_k(t-x_2)} \left(\int_0^{t-x_2} \dots \int_0^{t-x_2} 1(0 < x_3 < \dots < x_k) \times e^{(\theta_{k-1}-\theta_k)x_k} \dots e^{(\theta_2-\theta_3)x_3} dx_3 \dots dx_k \right) dx_2$$

Supongamos que (1.6.5) ocurre para alguna $n = j-1$,

$$\sum_{k=1}^n a_k(t) \prod_{i=1}^{k-1} (\theta_j - \theta_i) = a_1(t) + (\theta_j - \theta_1) \sum_{k=2}^j a_k(t) \prod_{i=2}^{k-1} (\theta_j - \theta_i)$$

$$\sum_{k=1}^n a_k(t) \prod_{i=1}^{k-1} (\theta_j - \theta_i) = a_1(t) + (\theta_j - \theta_1) \int_0^t e^{\theta_1 x_2} e^{\theta_j(t-x_2)} dx_2$$

$$\sum_{k=1}^n a_k(t) \prod_{i=1}^{k-1} (\theta_j - \theta_i) = e^{\theta_1 t} + e^{\theta_j t} (1 - e^{(\theta_1 - \theta_j)t}) = e^{\theta_j t}$$

Así, (1.6.5) ocurre para toda $n \geq 1$. En consecuencia, para $\mathbf{A} = \text{diag}(\boldsymbol{\theta})$ y $\theta_1, \dots, \theta_l$ distintos, (1.6.4) es cierta para todo $l \geq 1$. Supongamos ahora que \mathbf{A} es una matriz arbitraria con distintos valores propios $\theta_1, \dots, \theta_l$. Entonces no es difícil mostrar que

$$\exp(t\mathbf{A}) = \boldsymbol{\Phi} \text{diag}(e^{\theta_1 t}, \dots, e^{\theta_l t}) \boldsymbol{\Phi}^{-1}, \quad (1.6.6)$$

donde $\boldsymbol{\Phi} = (\phi_1^T, \dots, \phi_l^T)$ es una matriz de $l \times l$ que consiste de la columna derecha de los valores propios de \mathbf{A} . Entonces por la primera parte de la demostración tenemos que

$$\exp(t\mathbf{A}) = \boldsymbol{\Phi} \left(\sum_{k=1}^l a_k(t) \prod_{i=1}^{k-1} (\text{diag}(\boldsymbol{\theta}) - \theta_i \mathbf{I}) \right) \boldsymbol{\Phi}^{-1}$$

$$\exp(t\mathbf{A}) = \sum_{k=1}^l a_k(t) \prod_{i=1}^{k-1} (\boldsymbol{\Phi} \text{diag}(\boldsymbol{\theta}) \boldsymbol{\Phi}^{-1} - \theta_i \boldsymbol{\Phi} \boldsymbol{\Phi}^{-1})$$

$$\exp(t\mathbf{A}) = \sum_{k=1}^l a_k(t) \prod_{i=1}^{k-1} (\mathbf{A} - \theta_i \mathbf{I}).$$

Si no todos los valores propios de \mathbf{A} son distintos,

$$\mathbf{A} = \mathbf{C} \mathbf{D} \mathbf{C}^{-1},$$

donde \mathbf{D} es una matriz triangular superior y \mathbf{C} es una matriz no singular. Notemos que

$$\prod_{i=1}^l (\theta_i - s) = \det(\mathbf{A} - s\mathbf{I}) = \det(\mathbf{C}\mathbf{D}\mathbf{C}^{-1} - s\mathbf{C}\mathbf{C}^{-1})$$

$$\prod_{i=1}^l (\theta_i - s) = \det(\mathbf{D} - s\mathbf{I}) = \prod_{i=1}^l (d_{ii} - s).$$

Así, los valores propios de \mathbf{A} y \mathbf{D} coinciden, y es equivalente a $\{d_{11}, \dots, d_{ll}\} = \{\theta_1, \dots, \theta_l\}$.

Consideremos una matriz triangular $\mathbf{D}' = (d'_{ij})$ tal que $d'_{ij} = d_{ij}$ para $i \neq j$ y \mathbf{D}' tiene distintos elementos diagonales d'_{ii} con $|d'_{ii} - d_{ii}| < \varepsilon$ para toda $i = 1, \dots, l$ y alguna $\varepsilon > 0$. Entonces, la matriz $\mathbf{A}' = \mathbf{C}\mathbf{D}'\mathbf{C}^{-1}$ también tiene distintos valores propios y por (1.5.4)

$$\|\mathbf{A} - \mathbf{A}'\| = \|\mathbf{C}(\mathbf{D} - \mathbf{D}')\mathbf{C}^{-1}\| \leq l\|\mathbf{C}\|\|\mathbf{C}^{-1}\|\varepsilon$$

Por otra parte, no es difícil mostrar que $\|\exp(t\mathbf{A}') - \exp(t\mathbf{A})\| \rightarrow 0$ siempre que $\|\mathbf{A}' - \mathbf{A}\| \rightarrow 0$.

Si los valores propios $\theta_1, \dots, \theta_l$ de \mathbf{A} son distintos, entonces (1.6.6) inmediatamente implica que $\lim_{t \rightarrow \infty} e^{-st} \exp(t\mathbf{A}) = \mathbf{0}$ para cada $s > \max_{i \in E} \Re(\theta_i)$. De hecho, usando (1.6.6) tenemos que

$$e^{-st} \exp(t\mathbf{A}) = \Phi \text{diag}(e^{(\theta_1-s)t}, \dots, e^{(\theta_l-s)t}) \Phi^{-1} \rightarrow \mathbf{0}.$$

■

Lema 1.6.3: Si $\mathbf{A}(t)$ y $\mathbf{A}'(t)$ son dos funciones diferenciales matriciales, entonces

$$\frac{d}{dt}(\mathbf{A}(t)\mathbf{A}'(t)) = \left(\frac{d}{dt}\mathbf{A}(t)\right)\mathbf{A}'(t) + \mathbf{A}(t)\left(\frac{d}{dt}\mathbf{A}'(t)\right).$$

Lema 1.5.4: Si \mathbf{A} es no singular, entonces

$$\int_v^t \exp(x\mathbf{A})dx = \mathbf{A}^{-1}(\exp(t\mathbf{A}) - \exp(v\mathbf{A})).$$

Si todos los valores propios de \mathbf{A} tienen parte real negativa, entonces

$$\int_0^{\infty} \exp(x\mathbf{A})dx = -\mathbf{A}^{-1}.$$

Demostración: Sea la matriz de funciones $\mathbf{F}(t)$ diferenciable. Entonces

$$\int_v^t \mathbf{F}^{(1)}(x)dx = \mathbf{F}(t) - \mathbf{F}(v) \text{ para } v < t. \text{ Sea } \mathbf{F}(x) = \mathbf{A}^{-1} \exp(x\mathbf{A})$$

$$\left(\frac{d}{dx}\right)\mathbf{A}^{-1} \exp(x\mathbf{A}) = \mathbf{A}^{-1}\left(\frac{d}{dx}\right)\exp(x\mathbf{A}) = \mathbf{A}^{-1}\mathbf{A} \exp(x\mathbf{A}) = \exp(x\mathbf{A}).$$

■

Capítulo 2: Procesos de Markov

2.1 Introducción

A veces es conveniente tener un modelo que describa situaciones donde los estados cambian en puntos del tiempo arbitrarios. Esto se logra considerando una colección de variables aleatorias $\{X(t), t \geq 0\}$ donde $t \in \mathbb{R}^+$. Un proceso de Markov es un proceso estocástico $\{X(t), t \geq 0\}$ donde $t \in \mathbb{R}^+$ con un espacio de estados discreto o continuo, tal que

$$\mathbb{P}\left(X_{t_n} = i_n \mid X_{t_{n-1}} = i_{n-1}, \dots, X_{t_1} = i_1, X(0) = i_0\right) = \mathbb{P}\left(X_{t_n} = i_n \mid X_{t_{n-1}} = i_{n-1}\right),$$

donde \mathbb{P} denota la probabilidad de ocurrencia del evento (Bladt, 2004).

Sea E un espacio de estados discreto (finito o contable) y $\{X(t), t \geq 0\}$ un proceso de Markov con espacio de estados E con semigrupo de transición $\{P\}_{t \geq 0}$. Escribimos $p_{ij}^t = P^t(i, j) = \mathbb{P}_i(X_t = j)$ donde $\mathbb{P}_i(\cdot)$ denota la probabilidad de que el proceso de Markov $\{X(t), t \geq 0\}$ pase del estado i a cualquier otro estado del espacio de estados E al tiempo $t \geq 0$, y podemos identificar el semigrupo de transición por la familia $P^t_{t \geq 0} = (p_{ij}^t)_{t \geq 0}$ de matrices de transición. Las ecuaciones de Chapman-Kolmogorov $P^{t+s} = P^t P^s$ pueden ser interpretadas en el sentido usual de la multiplicación matricial. (Asmussen.2000)

Una estructura de salto puro puede ser interpretada como el tiempo de estancia en cada estado donde esta sea positiva para que las trayectorias muestrales sean constantes por pedazos. Para un proceso de saltos puro, denotamos el tiempo de saltos por $S_0 = 0 < S_1 < S_2 < \dots$, los tiempos de espera por $T_n = S_{n+1} - S_n$ y la secuencia de estados visitados por Y_0, Y_1, \dots . Así las trayectorias muestrales son constantes entre tiempos de saltos consecutivos S_n y definimos el valor en S_n continuo por la derecha, i.e. $X_{S_n} = Y_n$. El proceso puede ser absorbido en i . En ese caso hay un último S_n (tiempo de absorción) finito y se usa la convención $T_n = T_{n+1} = \dots = \infty$, $Y_n = Y_{n+1} = \dots = i$. Esto es una estructura muy simple de trayectorias. Un problema desde este punto de vista es la posibilidad de que el tiempo de explosión $\omega(\Delta) = \sup_n S_n$ sea finito. Esto parece contrario a la intuición en muchos casos, pero es perfectamente posible desde el punto de vista de la teoría general

Teorema 2.1.1: Un proceso de Markov $\{X_t\}_{t \geq 0}$ tiene la propiedad fuerte de Markov i.e.

$$\mathbb{P}\left(X_{t+h_1} = i_1, \dots, X_{t+h_n} = i_n \mid \mathcal{F}_\tau\right) = \mathbb{P}\left(X_{t+h_1} = i_1, \dots, X_{t+h_n} = i_n \mid X_0 = \sigma(X_\tau)\right)$$

con respecto a cualquier tiempo de paro τ el cual asume solamente un número contable de valores, $\tau \in \{\infty, s_1, s_2, \dots\}$.

Definición 2.1.1: Sea $\{X(t), t \geq 0\}$ un proceso estocástico en tiempo discreto definido sobre un espacio de probabilidad (Ω, F, \mathbb{P}) y tomado valores en un conjunto discreto E . $\{X(t)\}$ es una cadena de Markov si y solo si

$$\mathbb{P}(X_{t+m} = i_{t+m}, \dots, X_{t+1} = i_{t+1} | X_t = i_t, \dots, X_0 = i_0) = \mathbb{P}(X_{t+m} = i_{t+m}, \dots, X_{t+1} = i_{t+1} | X_t = i_t)$$

para toda $t, m \in \mathbb{N}$, $i_1, \dots, i_{t+m} \in E$. Además $\{X(t), t \geq 0\}$ es homogéneo en tiempo si $\mathbb{P}(X_{t+1} = j | X_t = i)$ no depende de t . En caso afirmativo se denota dicha probabilidad p_{ij} y se interpreta como la probabilidad de transición del estado i al estado j .

El siguiente resultado describe la estructura básica de un proceso de Markov de saltos hasta el tiempo de explosión.

Teorema 2.1.2: Consideremos un proceso de Markov de saltos. Entonces la distribución conjunta de las secuencias $\{Y_n\}$ de estados visitados y de los $\{T_n\}$ tiempos de espera está dada por: (i) $\{Y_n\}$ es una cadena de Markov; (ii) existe $\lambda(i) \geq 0$ tal que, dada $\{Y_n\}$, el T_i es independiente, con T_k siendo distribuida exponencialmente con intensidad $\lambda(Y_k)$.

Demostración: La distribución conjunta de Y_n y T_n está especificada completamente por probabilidades de la forma $\mathbb{P}_i F_n$ donde $F_n = \{Y_k = i(k), T_{k-1} > t(k), k = 1, \dots, n\}$. Denotemos a $i(0) = i$. El teorema es equivalente a

$$\mathbb{P}_i F_n = \prod_{k=1}^n q_{i(k-1)i(k)} \exp\{-\lambda(i(k-1))t(k)\} \quad (2.1.1)$$

para alguna matriz de transición \mathbf{Q} e intensidades convenientes $\lambda(i)$. Es claro que el único candidato posible para \mathbf{Q} es $q_{ij} = \mathbb{P}_i(Y_1 = j)$. Para determinar $\lambda(i)$, sea $z(t) = \mathbb{P}_i(T_0 > t)$. Entonces $X_t = i$ en $\{T_0 > t\}$, la propiedad de Markov nos conduce a

$$z(t+s) = \mathbb{E}_i \mathbb{P}_i(T_0 > t+s | \sigma(X_s; s \leq t)) = \mathbb{E}_i [\mathbb{P}_{j_i}(T_0 > s); T_0 > t] z(s) z(t)$$

Entonces z es no creciente con $z(0) = 1$, hechos elementales en las ecuaciones fundamentales nos conduce a $z(t) = e^{-\lambda(i)t}$ para alguna $\lambda(i) \geq 0$.

Aplicando la propiedad de Markov una vez más obtenemos

$$\mathbb{P}_i(Y_1 = j, T_0 > t) = \mathbb{E}_i \mathbb{P}_i(Y_1 = j, T_0 > t | F_t)$$

$$\mathbb{P}_i(Y_1 = j, T_0 > t) = \mathbb{E}_i [\mathbb{P}_i(Y_1 = j); T_0 > t] = q_{ij} e^{-\lambda(i)t}$$

la cual es (2.1.1.). El caso para $n > 1$ se sigue por inducción y la propiedad fuerte de Markov. Evaluando $\mathbb{P}_i F_n$ condicionando sobre $\sigma(X_s; s \leq S_{n-1})$ obtenemos de $X_{S_{n-1}} = Y_{n-1}$ que

$$\mathbb{P}_i F_n = \mathbb{E}_i \left[\mathbb{P}_{X_{S_{n-1}}} (Y_1 = i(n), T_0 > t(n)); F_{n-1} \right] = q_{ij} e^{-\lambda(i)t}$$

$$\mathbb{P}_i F_n = \mathbb{P}_{i(n-1)} (Y_1 = i(n), T_0 > t(n)) \mathbb{P}_i (F_{n-1})$$

$$\mathbb{P}_i F_n = q_{i(n-1)i(n)} e^{-\lambda(i(n-1))t(n)} \mathbb{P}_i (F_{n-1}).$$

■

2.2 Construcción minimal

La descripción intuitiva de un modelo práctico es usualmente dada en términos de las intensidades $\lambda(i)$ y las probabilidades de salto q_{ij} en lugar de en términos de las matrices de transición \mathbf{P}^t las cuales son difíciles de evaluar. Por lo tanto surge la pregunta si cualquier conjunto de $\lambda(i)$, q_{ij} nos lleva a un proceso de Markov de saltos.

La construcción es inmediatamente sugerida por el Teorema 2.1.2.

Suponemos por lo tanto que dado un conjunto $\lambda(i) > 0$, $i \in E$, y una matriz de transición \mathbf{Q} en E (la matriz \mathbf{Q} del Teorema 2.1.2. tiene la propiedad $q_{ii} = 0$ si y solo si $\lambda(i) > 0$, pero esto no es supuesto aquí). Sea $\Delta \notin E$ algún estado extra (necesario para describir el proceso después de una posible explosión), definimos $E_\Delta = E \cup \{\Delta\}$ y $\lambda(\Delta) = 0$, $q_{\Delta\Delta} = 1$. Se considera el espacio muestral

$$\Omega = (0, \infty]^{\mathbb{N}} \times E_\Delta^{\mathbb{N}} = \{(t_0, t_1, \dots, y_0, y_1, \dots) : 0 < t_k \leq \infty, y_k \in E_\Delta\}$$

y sean $T_0, T_1, \dots, Y_0, Y_1, \dots$ las proyecciones en Ω . Las probabilidades en Ω tienen las siguientes propiedades:

1. $\{Y_n\}$ es una cadena de Markov con matriz de transición \mathbf{Q} y $\mathbb{P}_i(Y_0 = i) = 1$;
2. dados $\{Y_n\}$, los T_l son independientes, con T_k distribuida exponencialmente con intensidad $\lambda(i)$ en $\{Y_k = i\}$.

Construimos $\{X_t\}_{t \geq 0}$ hasta el tiempo de la explosión simplemente invirtiendo la construcción de los Y_k , sea $S_0 = 0$,

$$S_n = T_0 + \dots + T_{n-1}, \quad \omega(\Delta) = \sup S_n = T_0 + T_1 + \dots,$$

$$X_t = \begin{cases} Y_k & \text{si } S_k \leq t < S_{k+1} \\ \Delta & \text{si } t \geq \omega(\Delta) \end{cases}$$

Teorema 2.2.1: $\{X_t\}_{t \geq 0}$ es un proceso de Markov de saltos en E_Δ .

Necesitamos estudiar el tiempo residual de estancia R_t , al tiempo t , i.e. $R_t = S_{n(t)} - t$, donde $n(t) = \min\{n : S_n > t\}$.

Lema 2.2.1: Dado $\sigma(X_s; s \leq t)$, la distribución de R_t es exponencial con intensidad $\lambda(X_t)$.

Demostración: Dado $\sigma(X_s; s \leq t)$, la distribución de $T_{n(t)-1}$ de T dado $T > u$, donde $u = t - S_{n(t)-1}$ y T es exponencial con intensidad $\lambda(Y_{n(t)-1}) = \lambda(X_t)$. Debemos demostrar que

$$\mathbb{P}_i(R_t > r, A) = \mathbb{E}_i[\exp\{-\lambda(X_t)r\}; A] \quad (2.2.1)$$

para toda $r < \infty$ y toda $A \in \sigma(X_s; s \leq t)$. Si $A \subseteq \{\omega(\Delta) \leq t\}$, entonces ambos lados son $\mathbb{P}_i A$ de modo que podemos asumir $A \subseteq \{\omega(\Delta) > t\}$ y entonces es suficiente considerar a A de la forma $\{n(t) - 1 = n, F_n\} = \{F_n, S_n \leq t, S_n + T_n > t\}$. Así, si condicionamos sobre las $Y_k, T_{k-1}, k = 1, \dots, n$ y usamos la fórmula $\mathbb{P}(T > t + r) = e^{-\lambda r} \mathbb{P}(T > t)$ para la distribución exponencial, podemos evaluar $\mathbb{P}_i(R_t > r, A)$ como

$$\mathbb{P}_i(F_n, S_n \leq t, S_n + T_n > t + r) = e^{-\lambda(i(n))r} \mathbb{P}_i(F_n, S_n \leq t, S_n + T_n > t).$$

■

Demostración del Teorema 2.2.1: $\{X_t\}_{t \geq 0}$ es un salto puro, de modo que es suficiente demostrar que en $\{X_t = i\}$ la distribución condicional de $\{X_{t+s}\}_{s \geq 0}$ dado $\sigma(X_s; s \leq t)$ es la \mathbb{P}_i distribución de $\{X_{t+s}\}_{s \geq 0}$. Definimos

$$M_t = (T_{n(t)}, T_{n(t)+1}, \dots, Y_{n(t)-1}, Y_{n(t)}, \dots).$$

Entonces $\{X_{t+s}\}_{s \geq 0}$ es construida de (R_t, M_t) de la misma forma que $\{X_t\}_{t \geq 0}$ es construida de

$$(R_0, M_0) = (T_0, M_0) = (T_0, T_1, \dots, Y_0, Y_1, \dots).$$

Por lo tanto mostraremos que en $\{X_t = i\}$, la distribución condicional de (R_t, M_t) dado $\sigma(X_s; s \leq t)$ es la \mathbb{P}_i distribución de (R_0, M_0) en la distribución condicional, R_t, M_t son independientes, R_t tiene la \mathbb{P}_i distribución de R_0 , M_t tiene la \mathbb{P}_i distribución de M_0 .

Ahora $\{(Y_n, T_n)\}$ es una cadena de Markov con espacio de estados $E_\Delta \times (0, \infty)$ y kernel de transición dado por

$$\mathbb{P}(Y_{n+1} = j, T_{n+1} > t \mid \sigma(Y_k, T_k : k \leq n)) = q_{Y_n j} e^{-\lambda(j)t} \quad (2.2.2)$$

También $n(t) - 1$ es un tiempo de paro con respecto a la cadena y debemos evaluar la distribución de (R_t, M_t) condicionalmente sobre $\sigma(X_s; s \leq t)$ primero condicionando sobre la σ -álgebra $\sigma(Y_k, T_k : k \leq n(t) - 1)$ entonces $Y_{n(t)-1} = X_t$, la propiedad fuerte de Markov y (2.2.2) implican que dado $\sigma(Y_k, T_k : k \leq n(t) - 1)$, M_t tiene la \mathbb{P}_{X_t} distribución de M_0 , mientras que R_t es degenerado. Estos hechos y la medibilidad de X_t implican que R_t, M_t son independientes y que M_t tiene la \mathbb{P}_i distribución de M_0 , mientras que R_t tiene la \mathbb{P}_i distribución de R_0 por el Lema (2.2.1).

Proposición 2.2.1: Definimos $R = \sum_{i=0}^{\infty} \lambda(Y_i)^{-1}$. Entonces para cualquiera $i \in E$, los conjuntos $\{\omega(\Delta) < \infty\}$ y $\{R < \infty\}$ coinciden \mathbb{P}_i casi seguramente.

Demostración: Condicionando sobre $\{Y_n\}$, $\omega(\Delta) = \sup S_n = \sum_{i=0}^{\infty} T_i$ se distribuye como

$\sum_{i=0}^{\infty} \lambda(Y_i)^{-1} V_i$, donde las V_i son variables aleatorias exponenciales independientes idénticamente distribuidas con parámetro 1. El resultado por lo tanto proviene por hechos estándar en las sumas ponderadas de variables aleatorias independientes idénticamente distribuidas. Así $R < \infty$ implica que $\omega(\Delta) < \infty$ porque de $R = \mathbb{E}(\omega(\Delta) | Y_0, Y_1, \dots)$, y la implicación inversa puede ser vista, aplicando el criterio de las tres series. ■

Proposición 2.2.2: Criterios suficientes para $\mathbb{P}_i(\omega(\Delta) < \infty) = 0$ para toda $i \in E$ son los siguientes: $\sup_{i \in E} \lambda(i) < \infty$; E es finito; $\{Y_n\}$ es recurrente.

Demostración: Se sigue de la Proposición 2.2.1 que $\lambda(Y_n) \rightarrow \infty$ en $\{\omega(\Delta) < \infty\}$. Por lo tanto la suficiencia de $\sup_{i \in E} \lambda(i) < \infty$ es clara, y que E sea finito es consecuencia de que el $\sup_{i \in E} \lambda(i) < \infty$. Si $\{Y_n\}$ es recurrente, y $X_0 = Y_0 = i$, entonces $\lambda(i)$ es un punto límite de $\{\lambda(Y_n)\}$. Así $\lambda(Y_n) \rightarrow \infty$ no puede ocurrir, para que $R = \infty$ y $\mathbb{P}_i(\omega(\Delta) < \infty) = 0$. ■

2.3 La matriz de intensidad

2.3.1 Definición y Unicidad

Ahora supongamos que $q_{ii} = 0$ cuando $\lambda(i) > 0$ y definimos la matriz de intensidad $\Lambda = (\lambda(i, j))_{i, j \in E}$ del proceso por

$$\lambda(i, j) = \lambda(i)q_{ij}, \quad j \neq i, \quad \lambda(i, i) = -\lambda(i) \quad (2.3.1.1)$$

Proposición 2.3.1: Una matriz Λ de $E \times E$ es la matriz de intensidad de un proceso de Markov de saltos $\{X_t\}$ si y solo si

$$\lambda(i, i) \leq 0, \lambda(i, j) \geq 0, j \neq i, \sum_{j \in E} \lambda(i, j) = 0. \quad (2.3.1.2)$$

Además, Λ está en correspondencia uno a uno con la distribución del proceso minimal.

Demostración: Si Λ es una matriz de intensidad, se sigue de (2.3.1.1) considerando los casos $\lambda(i) = 0$ y $\lambda(i) > 0$ separadamente tal que $\sum_{j \neq i} \lambda(i, j) = \lambda(i)$ y por lo tanto (2.3.1.2) ocurre.

Inversamente si (2.3.1.2) se satisface, entonces sea $\lambda(i) = -\lambda(i, i)$, definimos q_{ij} por (2.3.1.1) y $q_{ii} = 0$ si $\lambda(i) > 0$, y sea $q_{ij} = \delta_{ij}$ en otro caso. Entonces \mathbf{Q} es la matriz de transición, y claramente el proceso de Markov de saltos determinado por \mathbf{Q} y los $\lambda(i)$ tienen la matriz de intensidad Λ . ■

2.3.2 Ecuaciones Diferenciales de Kolmogorov

Teorema 2.3.2.1: Sea Λ una matriz de intensidad en E y $\{X_t\}$ el proceso de Markov de saltos minimal correspondiente en E , $p_{ij}^t = \mathbb{P}_i(X_t = j)$. Entonces las matrices de $E \times E$ $\mathbf{P}^t = (p_{ij}^t)$ satisfacen la ecuación $\left(\frac{d}{dt}\right)\mathbf{P}^t = \Lambda\mathbf{P}^t$, i.e.

$$\frac{dp_{ij}^t}{dt} = \sum_{k \in E} \lambda(i, k) p_{kj}^t \quad (2.3.2.1)$$

y la ecuación $\left(\frac{d}{dt}\right)\mathbf{P}^t = \mathbf{P}^t\Lambda$, i.e.

$$\frac{dp_{ij}^t}{dt} = \sum_{k \in E} p_{ik}^t \lambda(k, j). \quad (2.3.2.2)$$

Demostración: Condicionando sobre $T_0 = s$

$$p_{ij}^t = \mathbb{P}_i(T_0 > t) \delta_{ij} + \int_0^t \lambda(i) e^{-\lambda(i)s} \sum_{k \neq i} q_{ik} p_{kj}^{t-s} ds$$

$$p_{ij}^t = e^{-\lambda(i)t} \left[\delta_{ij} + \int_0^t \sum_{k \neq i} \lambda(i, k) e^{\lambda(i)s} p_{kj}^s ds \right].$$

El integrando $f(s) = \sum_{k \neq i} \lambda(i, k) e^{\lambda(i)s} p_{kj}^s$ está bien definido con $\sup_{s \leq T} f(s)$ para toda $T < \infty$ entonces $\sum |\lambda(i, k)| = 2\lambda(i) < \infty$. Esto muestra primero que p_{ij}^t es continua y a partir de esto $f(s)$ es continua. Por lo tanto p_{ij}^t es diferenciable con derivada

$$-\lambda(i) e^{-\lambda(i)t} \left[\delta_{ij} + \int_0^t f(s) ds \right] + e^{-\lambda(i)t} f(t) = -\lambda(i) p_{ij}^t + \sum_{k \neq i} \lambda(i, k) p_{kj}^t = \sum_{k \in E} \lambda(i, k) p_{kj}^t.$$

La demostración de la ecuación $\left(\frac{d}{dt} \right) \mathbf{P}^t = \mathbf{P}^t \mathbf{\Lambda}$ solamente se dará bajo la suposición

$$\sup_{i \in E} \lambda(i) < \infty \quad (2.3.2.3)$$

la cual puede ser usada para inferir que $\frac{(p_{kj}^s - \delta_{kj})}{s}$ está uniformemente acotada en s, j, k . Se sigue entonces que

$$0 \leq p_{kj}^s \leq \lambda(k) \int_0^s e^{-\lambda(k)u} du, \quad k \neq j$$

$$0 \leq 1 - p_{kk}^s \leq \lambda(k) \int_0^s e^{-\lambda(k)u} du$$

y (2.3.2.2) se obtiene por convergencia dominada (usando $\sum p_{ik}^t < \infty$) de

$$\frac{p_{ij}^{t+s} - p_{ij}^t}{s} = \sum_{k \in E} p_{ik}^t \frac{p_{kj}^s - \delta_{kj}}{s} \rightarrow \sum_{k \in E} p_{ik}^t \lambda(k, j).$$

■

En el caso de un espacio de estados finito E , los resultados sobre existencia y unicidad de los sistemas de ecuaciones diferenciales lineales junto con $\mathbf{P}^0 = \mathbf{I}$ nos lleva al siguiente corolario.

Corolario 2.3.2.1: Si E es finito, entonces $\mathbf{P}^t = e^{\mathbf{\Lambda}t} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{\Lambda}^n$, $t > 0$.

2.4 Estacionalidad y resultados límite

2.4.1 Clasificación de Estados

Proposición 2.4.1.1: Las siguientes propiedades son equivalentes:

- (a) $\{Y_n\}$ es irreducible;
- (b) para cualquier $i, j \in E$ tenemos que $p_{ij}^t > 0$ para alguna $t > 0$;
- (c) para cualquier $i, j \in E$ tenemos que $p_{ij}^t > 0$ para toda $t > 0$.

Demostración: Denotemos por

$$\omega(i) = \inf\{t > 0 : X_t = i, \lim_{s \uparrow t} X_s \neq i\} \quad (2.4.1.1)$$

al tiempo de llegada a i si $X_0 \neq i$ y el tiempo de recurrencia de i si $X_0 = i$. Entonces j tiene un tiempo de arribo exponencial. Es claro que $p_{ij}^t > 0$ si y sólo si $\mathbb{P}_i(\omega(j) \leq t) > 0$. Ahora $\mathbb{P}_i(\omega(j) \leq t) > 0$ si y solo si alguna trayectoria $i i_1 \dots i_n j$ de i a j es posible para $\{Y_n\}$, y en ese caso podemos evaluar la distribución condicional F de $\omega(j)$ dado $\omega(j) < \infty$ condicionando en las trayectorias. Así F es una mezcla de convoluciones de distribuciones exponenciales con intensidades $\lambda(i, i_1), \lambda(i_1, i_2), \dots$ y por lo tanto tiene una densidad mayor que cero en $(0, \infty)$. Así $\mathbb{P}_i(\omega(j) \leq t) > 0$ si y solo si $\{Y_n\}$ puede llegar a j desde i . ■

2.4.2 Medidas Estacionarias

Una medida $\mathbf{v} \neq \mathbf{0}$ es estacionaria si $0 \leq v_j < \infty$ y $\mathbf{vP}^t = \mathbf{v}$ para toda t .

Teorema 2.4.2.1: Supongamos que $\{X_t\}$ es irreducible y recurrente en E . Entonces existe una medida estacionaria \mathbf{v} . Esta \mathbf{v} tiene la propiedad $v_j > 0$ para toda j y puede ser encontrada de una de las siguientes maneras:

- (i) para algún estado fijo y arbitrario i , v_j es el tiempo transcurrido esperado entre entradas exitosas de j a i . Esto es, con $\omega(i)$ dado por

$$v_j = \mathbb{E}_i \int_0^{\omega(i)} I(X_t = j) dt \quad (2.4.2.1)$$

- (ii) $v_j = \frac{\mu_j}{\lambda(j)}$, donde $\boldsymbol{\mu}$ es estacionario para $\{Y_n\}$;

- (iii) es solución de $\mathbf{v}\boldsymbol{\Lambda} = \mathbf{0}$

Demostración: Primero probaremos la unicidad considerando la cadena de Markov X_0, X_1, \dots . Esta es irreducible, entonces todas las entradas $p_{ij}^t > 0$, y cualquier \mathbf{v} estacionario para $\{X_t\}$ es también estacionario para $\{X_n\}$, entonces si la cadena es irreducible y recurrente existe una medida estacionaria \mathbf{v} , que satisface $v_j > 0$ para toda j y es una constante multiplicativa única, entonces tenemos que demostrar que $\{X_n\}$ es recurrente. Pero para cualquier i , la secuencia U_1, U_2, \dots de tiempos de espera de i es infinita, entonces i es recurrente. Las U_k son independientes idénticamente distribuidas con $\mathbb{P}(U_k > 1) > 0$ tenemos que $U_k > 1$ y por lo tanto también $X_n = i$.

Para (i) mostramos la estacionalidad de (2.4.2.1) evaluando el j -ésimo componente de \mathbf{vP}^h . Primero notemos que $\{X_t\}_{0 \leq t \leq h}$ y $\{X_{\omega(i)+t}\}_{0 \leq t \leq h}$ tiene la misma \mathbb{P}_i distribución porque $X_{\omega(i)} = i$.

Por lo tanto

$$v_j = \mathbb{E}_i \left[\int_0^h + \int_h^{\omega(i)} I(X_t = j) dt \right] = \mathbb{E}_i \left[\int_{\omega(i)}^{\omega(i)+h} + \int_h^{\omega(i)} I(X_t = j) dt \right] = \mathbb{E}_i \left[\int_h^{\omega(i)+h} I(X_t = j) dt \right]$$

$$v_j = \mathbb{E}_i \int_0^{\omega(i)} I(X_{t+h} = j) dt$$

notemos que la primera igualdad es válida también si $\omega(i) < h$. Así

$$v_j = \mathbb{E}_i \int_0^{\infty} \mathbb{P}(X_{t+h} = j, \omega(i) > t | F_t) dt = \mathbb{E}_i \int_0^{\infty} p_{X,h}^h I(\omega(i) > t) dt$$

$$v_j = \sum_{k \in E} p_{kj}^h \mathbb{E}_i \int_0^{\infty} I(X_t = k, \omega(i) > t) dt = \sum_{k \in E} v_k p_{kj}^h$$

demostrando $\mathbf{vP}^h = \mathbf{v}$ y (i). Con $\tau(i) = \inf\{n : Y_n = i\}$, tenemos que

$$v_j = \mathbb{E}_i \int_0^{\omega(i)} I(X_t = j) dt = \mathbb{E}_i \sum_{n=0}^{\tau(i)-1} T_n I(Y_n = j)$$

$$v_j = \sum_{n=0}^{\infty} \mathbb{E}_i \mathbb{E}_i [T_n; Y_n = j, \tau(i) > n | \{Y_n\}_0^{\infty}]$$

$$v_j = \frac{1}{\lambda(j)} \mathbb{E}_i \sum_{n=0}^{\infty} I(Y_n = j, \tau(i) > n) = \frac{1}{\lambda(j)} \frac{\mu_j}{\mu_i}$$

Esto es, v_j es proporcional a $\frac{\mu_j}{\lambda(j)}$, demostrando (ii)

Para (iii), notemos que de acuerdo con (ii) \mathbf{v} es estacionario para $\{X_t\}$ si y solo si $(v_j \lambda(j))_{j \in E}$ es estacionario para $\{Y_n\}$, i.e. si y solo si

$$\sum_{i \in E} v_i \lambda(i) q_{ij} = v_j \lambda(j) \text{ para toda } j \in E,$$

entonces $q_{ii} = 0$, si y solo si

$$0 = -v_j \lambda(j) + \sum_{i \neq j} v_i \lambda(i, j) = \sum_{i \in E} v_i \lambda(i, j).$$

Finalmente, $0 < v_j < \infty$ se sigue fácilmente por (ii), entonces en el caso recurrente $0 < \lambda(j) < \infty$ ■

En este capítulo se mostraron los procesos de Markov, sus propiedades, las ecuaciones diferenciales de Kolmogorov para los procesos de Markov; por último se describió la clasificación de estados y las medidas estacionarias.

Capítulo 3: Distribuciones Tipo Fase

3.1 Introducción

Las distribuciones tipo fase están basadas en el método de estados, una técnica introducida por Erlang (1917) y generalizada por Neuts (1975). La idea principal es modelar intervalos aleatorios de tiempo como un número de segmentos distribuidos exponencialmente y aprovechar la estructura Markoviana para simplificar el análisis.

Erlang definió una variable aleatoria no negativa como el tiempo que toma moverse a través de un número fijo de estados, utilizando una cantidad de tiempo exponencial con una tasa fija positiva cada una hasta la absorción.

La distribución de una variable aleatoria es definida a través de una matriz; su función de densidad, momentos, etc. están expresados en términos de esta matriz.

Las distribuciones tipo fase proporcionan una estructura simple para demostrar cómo se pueden extender resultados sobre distribuciones exponenciales a modelos más complejos sin perder la manejabilidad computacional.

Las distribuciones tipo fase han sido usadas en una amplia gama de aplicaciones de modelación estocástica en áreas tan diversas como las telecomunicaciones, modelación del tráfico, bioestadística, teoría de colas, cinética de los medicamentos, teoría de la credibilidad, análisis de supervivencia y teoría del riesgo (Fackrell. 2003)

Las distribuciones tipo fase constituye una clase de distribuciones muy versátil que están definidas en los números reales no negativos; eso lleva a modelos que son algorítmicamente tratables. Su formulación también permite la estructura de Markov de modelos estocásticos cuando reemplazan a la distribución exponencial.

Si un problema puede ser resuelto explícitamente cuando las distribuciones son exponenciales, entonces el problema puede admitir una solución algorítmica involucrando un razonable grado de esfuerzo computacional si uno permite la suposición más general de estructura tipo fase.

3.2 Definición y Propiedades Básicas de las Distribuciones Tipo Fase

Consideremos un proceso de Markov con espacio de estados $E = \{1, \dots, m\}$. Supongamos que ninguno de los estados $i \in E = \{1, \dots, m\}$ del proceso de Markov es absorbente (Neuts. 1981).

Extenderemos el espacio de estados E agregando un estado que denotaremos como $m+1$ y el cual supondremos que es absorbente. Para el espacio de estados extendido $E' = \{1, \dots, m+1\}$, consideraremos la matriz de intensidad de un proceso de Markov no terminal $\mathbf{Q} = (q_{ij})_{i,j \in E'}$ escrita por bloques de la siguiente forma:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix}$$

donde $\mathbf{t} = (t_1 \ \dots \ t_m)'$ es un vector columna de dimensión l con componentes no negativos t_i , $\mathbf{T} = (t_{ij})$ es una matriz de $m \times m$ con $t_{ij} \geq 0$ para $i \neq j$ y $t_{ii} \leq 0$ tal que $\mathbf{t} + \mathbf{T}\mathbf{e} = \mathbf{0}$, entonces $\mathbf{t} = -\mathbf{T}\mathbf{e}$. Donde \mathbf{e} es un vector columna con todos sus componentes iguales a uno.

La interpretación del vector columna \mathbf{t} es el vector de la tasa de salida, i.e. el i -ésimo componente t_i brinda la intensidad en el estado i para dejar E

Para asegurar la absorción en tiempo finito con probabilidad uno, se asume que cada estado no absorbente es transitorio.

Una condición equivalente está dada por el siguiente lema.

Lema 3.2.1: Los estados $1, \dots, m$ son transitorios si y solo si la matriz \mathbf{T} es no singular.

Demostración: Sea a_i , $1 \leq i \leq m$ la probabilidad de absorción en el estado $m+1$, iniciando en el estado i . Las probabilidades a_i satisfacen el sistema de ecuaciones lineales

$$a_i = (-t_i T_{ii}^{-1}) + \sum_{v \neq i} (-T_{iv} T_{ii}^{-1}) a_v, \text{ para } 1 \leq i \leq m$$

o equivalentemente, $\mathbf{t} + \mathbf{T}\mathbf{a} = \mathbf{0}$.

Si \mathbf{T} es no singular, la solución $\mathbf{a} = \mathbf{e}$ es única, para que la absorción desde cualquier estado i sea forzosa.

Se muestra que si la absorción desde cualquier estado es forzosa, la matriz \mathbf{T} es no singular.

Si la matriz \mathbf{T} es singular, existe un vector positivo \mathbf{y} , tal que $\mathbf{y}\mathbf{T} = \mathbf{0}$. Esto implica que el producto punto $\mathbf{y}\exp(\mathbf{T}x)\mathbf{e} = \mathbf{y}\mathbf{e}$ es positivo para toda $x \geq 0$. Además, el vector $\mathbf{u} = \lim_{t \rightarrow \infty} \exp(\mathbf{T}t)\mathbf{e}$ no desaparece, entonces $\mathbf{y}\mathbf{u} = \mathbf{y}\mathbf{e} > 0$. Esto muestra que al menos hay un estado inicial $i \in \{1, \dots, m\}$ del cual la probabilidad $1 - u_i$ de una eventual absorción es menor que uno. Se sigue por contradicción que una absorción eventual desde cada estado inicial implica que \mathbf{T} es no singular. ■

La variable aleatoria η definida a continuación es llamada el tiempo de paro de $\{X_t\}$. Esta distribución es determinada por la distribución inicial $\boldsymbol{\alpha}' = (\alpha_1, \dots, \alpha_{m+1})$ de $\{X_t\}$ y por la matriz de subintensidades \mathbf{T} . Notemos que, en vez de $\boldsymbol{\alpha}'$, es suficiente considerar la función de probabilidad $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{m+1})$ en E .

Definición: La distribución del tiempo de paro hasta la absorción $\eta = \inf\{t \geq 0 \mid X_t = 0\}$ para un proceso de Markov finito de saltos $\{X_t\}$ con espacio de estados $E' = \{1, \dots, m+1\}$ es llamada distribución tipo fase con parámetros $(\boldsymbol{\alpha}, \mathbf{T})$. Denotaremos la distribución por $PH(\boldsymbol{\alpha}, \mathbf{T})$ (Bladt 2004).

Lema 3.2.2:

$$\exp(\mathbf{Q}x) = \begin{pmatrix} \exp(\mathbf{T}x) & \mathbf{e} - \exp(\mathbf{T}x)\mathbf{e} \\ \mathbf{0} & 1 \end{pmatrix}$$

Demostración: Para $n \geq 1$ tenemos que

$$\mathbf{Q}^n = \begin{pmatrix} \mathbf{T}^n & -\mathbf{T}^n \\ \mathbf{0} & 0 \end{pmatrix}$$

usando que $\mathbf{t} = -\mathbf{T}\mathbf{e}$. Entonces, si \mathbf{I}_n es la matriz identidad de dimensión $n \times n$, tenemos que

$$\exp(\mathbf{Q}x) = \sum_{n=0}^{\infty} \frac{\mathbf{Q}^n x^n}{n!}$$

$$\exp(\mathbf{Q}x) = \mathbf{I}_{p+1} + \sum_{n=1}^{\infty} \frac{\mathbf{Q}^n x^n}{n!}$$

$$\exp(\mathbf{Q}x) = \mathbf{I}_{p+1} + \sum_{n=1}^{\infty} \begin{pmatrix} \frac{(\mathbf{T}x)^n}{n!} & -\frac{(\mathbf{T}x)^n}{n!} \\ \mathbf{0} & 0 \end{pmatrix}$$

$$\exp(\mathbf{Q}x) = \mathbf{I}_{p+1} + \begin{pmatrix} \exp(\mathbf{T}x) - \mathbf{I}_p & -(\exp(\mathbf{T}x) - \mathbf{I}_p)\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix}$$

$$\exp(\mathbf{Q}x) = \begin{pmatrix} \exp(\mathbf{T}x) & \mathbf{e} - \exp(\mathbf{T}x)\mathbf{e} \\ \mathbf{0} & 1 \end{pmatrix}$$

■

Lema 3.2.3: La distribución $F(\cdot)$ del tiempo de absorción en el estado $m+1$, correspondiente al vector de probabilidades iniciales $(\alpha_1, \dots, \alpha_{m+1})$ está dada por

$$F(x) = 1 - \boldsymbol{\alpha} \exp(\mathbf{T}x)\mathbf{e}, \text{ para } x \geq 0. \quad (3.2.1)$$

Demostración: Las probabilidades $v_j(x)$, que el proceso este en el estado j al tiempo x , satisface el sistema de ecuaciones diferenciales

$$\mathbf{v}'(x) = \mathbf{v}(x)\mathbf{T}, \text{ para } x > 0$$

con condiciones iniciales $\mathbf{v}(0) = \boldsymbol{\alpha}$. La solución está dada por

$$\mathbf{v}(x) = \boldsymbol{\alpha} \exp(\mathbf{T}x).$$

Sustituyendo $\mathbf{v}(x)$ en la ecuación $F(x) = 1 - \mathbf{v}(x)\mathbf{e}$ tenemos que

$$F(x) = 1 - \boldsymbol{\alpha} \exp(\mathbf{T}x)\mathbf{e}.$$

■

Teorema 3.2.1: Consideremos el tiempo de absorción η con distribución tipo fase $PH(\boldsymbol{\alpha}, \mathbf{T})$. Entonces, para cada $t \geq 0$.

$$\mathbb{P}(\eta > t) = \boldsymbol{\alpha} \exp(t\mathbf{T})\mathbf{e}$$

Demostración: Se tiene que $\{\eta > t\} = \{X_t \neq 0\}$ si y solo si $X_t \in \{1, 2, \dots, p\}$ y consecuentemente por la ley de la probabilidad total

$$1 - F(t) = \mathbb{P}(\eta > t) = \mathbb{P}\boldsymbol{\alpha}(X_t \in \{0, 2, \dots, l-1\})$$

$$\mathbb{P}(\eta > t) = \sum_{j=0}^{l-1} \mathbb{P}\boldsymbol{\alpha}(X_t = j)$$

$$\mathbb{P}(\eta > t) = \sum_{j=0}^{l-1} \sum_{i=0}^{l-1} \mathbb{P}(X_t = j | X_0 = i) \mathbb{P}(X_0 = i)$$

$$\mathbb{P}(\eta > t) = (X(t) \neq l) = \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} \alpha_i p_{ij}(t)$$

$$\exp(t\mathbf{Q}) = \begin{pmatrix} \exp(t\mathbf{T}) & \mathbf{e} - \exp(t\mathbf{T})\mathbf{e} \\ \mathbf{0} & 1 \end{pmatrix}$$

■

Lema 3.2.4: Sea \mathbf{T} una matriz de subintensidad. Entonces $s\mathbf{I} - \mathbf{T}$ es no singular para cada $s \geq 0$ y todas las entradas de $(s\mathbf{I} - \mathbf{T})^{-1}$ son funciones racionales de $s \geq 0$. Además, para toda $s \geq 0$, $n \in \mathbb{N}$,

$$\int_0^{\infty} \exp(t(-s\mathbf{I} + \mathbf{T})) dt = (s\mathbf{I} - \mathbf{T})^{-1}$$

y

$$\frac{d^n}{ds^n} (s\mathbf{I} - \mathbf{T})^{-1} = (-1)^n n! (s\mathbf{I} - \mathbf{T})^{-n-1}$$

Demostración: Sea $s \geq 0$. Todos los valores propios de $\mathbf{T} - s\mathbf{I}$ tienen parte real negativa. Por lo tanto $s\mathbf{I} - \mathbf{T}$ es no singular. Demostraremos por inducción con respecto a n . Usando la regla de derivación tenemos que

$$\mathbf{0} = \frac{d}{ds} \mathbf{I} = \frac{d}{ds} ((s\mathbf{I} - \mathbf{T})(s\mathbf{I} - \mathbf{T})^{-1}) = (s\mathbf{I} - \mathbf{T})^{-1} + (s\mathbf{I} - \mathbf{T}) \frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{-1}$$

en consecuencia

$$\frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{-1} = -(s\mathbf{I} - \mathbf{T})^{-2}$$

Suponemos que es cierto para $n = 1, 2, \dots, k$. Entonces

$$\frac{d^{k-1}}{ds^{k-1}}(s\mathbf{I} - \mathbf{T})^{-1} = \frac{d}{ds} \left(\frac{d^k}{ds^k} (s\mathbf{I} - \mathbf{T})^{-1} \right) = (-1)^k k! \frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{-k-1}$$

$$\frac{d^{k-1}}{ds^{k-1}}(s\mathbf{I} - \mathbf{T})^{-1} = (-1)^{k+1} (k+1)! (s\mathbf{I} - \mathbf{T})^{-(k+1)-1}$$

porque

$$\mathbf{0} = \frac{d}{ds} \left((s\mathbf{I} - \mathbf{T})^{k+1} (s\mathbf{I} - \mathbf{T})^{-k-1} \right)$$

$$\mathbf{0} = (k+1)(s\mathbf{I} - \mathbf{T})^{-1} + (s\mathbf{I} - \mathbf{T})^{k+1} \frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{-k-1}.$$

■

Teorema 3.2.2: Supongamos que $\alpha_0 = 0$ y \mathbf{T} es una matriz de subintensidad no singular. Si $F \sim PH(\mathbf{a}, \mathbf{T})$, entonces F es continua con

Función de densidad $f(x) = \mathbf{a}e^{\mathbf{T}x}\mathbf{t}$, $t \geq 0$

Transformada de Laplace $\hat{l}(s) = \mathbf{a}(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{t}$

n-ésimo momento $\mu^{(n)} = (-1)^n n! \mathbf{a}\mathbf{T}^{-n}\mathbf{e}$

Demostración:

$$f(x) = -\frac{d}{dt} \mathbb{P}(\eta > x) = \mathbf{a}e^{\mathbf{T}x}(-\mathbf{T}\mathbf{e}t) = \mathbf{a}e^{\mathbf{T}x}\mathbf{t}$$

$$\hat{l}(s) = \int_0^\infty e^{-sx} f(x) dx = \int_0^\infty e^{-sx} \mathbf{a}e^{\mathbf{T}x}\mathbf{t} dx.$$

Puesto que $e^{-sx}\mathbf{I} = \exp(-sx\mathbf{I})$ entonces

$$\hat{l}(s) = \int_0^\infty \mathbf{a} \exp(-sx\mathbf{I}) \exp(\mathbf{T}x)\mathbf{t} dx.$$

Puesto que $\exp(-sx\mathbf{I})\exp(\mathbf{T}x) = \exp(x(-s\mathbf{I} + \mathbf{T}))$ entonces

$$\hat{l}(s) = \mathbf{a} \int_0^\infty \exp(x(-s\mathbf{I} + \mathbf{T})) dx \mathbf{t}$$

$$\hat{l}(s) = \mathbf{a}(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{t}.$$

Tomamos la n -ésima derivada de la transformada de Laplace. Entonces

$$\hat{l}^{(n)}(s) = \frac{d^n}{ds^n} \boldsymbol{\alpha}(s\mathbf{I} - \mathbf{T})^{-1} \mathbf{t} = (-1)^n n! \boldsymbol{\alpha}(s\mathbf{I} - \mathbf{T})^{-n-1} \mathbf{t}$$

Sustituyendo $s = 0$ obtenemos que $\mathbb{E}(X^n) = \mu^{(n)} = (-1)^n \hat{l}^{(n)}(0) = (-1)^n n! \boldsymbol{\alpha} \mathbf{T}^n \mathbf{e}$. ■

Un requerimiento adicional en la representación $PH(\boldsymbol{\alpha}, \mathbf{T})$ es que no existan estados superfluos. Las condiciones para que no existan estos estados se dan a conocer en la siguiente sección.

3.3 Representaciones Irreducibles

Se supone que en el estado $m+1$, se realizan instantáneamente ensayos multinomiales independientes con probabilidades $\alpha_1, \dots, \alpha_m, \alpha_{m+1}$ hasta que una de las m alternativas ocurra. Reiniciando el proceso \mathbf{Q} en el estado correspondiente, se considera el tiempo de la siguiente absorción y se repite el mismo procedimiento. Se observa que continuando el procedimiento indefinidamente se construye un nuevo proceso de Markov en el cual el estado $m+1$ es un estado instantáneo. Cuando el proceso entra al estado instantáneo, la probabilidad de que permanezca ahí para transiciones $r \geq 1$ está dada por $(1 - \alpha_{m+1}) \alpha_{m+1}^{r-1}$, y los números de visitas al estado instantáneo son independientes con una distribución geométrica común.

Si además consideramos la versión del proceso de Markov en el cual las funciones de la trayectoria son continuas por la derecha, se obtiene un proceso de Markov en $\{1, \dots, m\}$ con generador infinitesimal

$$\mathbf{Q}^* = \mathbf{T} + \mathbf{T}^0 \mathbf{A}^0 \quad (3.3.1)$$

donde \mathbf{T}^0 es una matriz de $m \times m$ con columnas idénticas \mathbf{T}^0 y $\mathbf{A}^0 = (1 - \alpha_{m+1})^{-1} \text{diag}(\alpha_1, \dots, \alpha_{m+1})$. La matriz $\mathbf{T}^0 \mathbf{A}^0$ puede ser escrita como

$$\mathbf{T}^0 \mathbf{A}^0 = (1 - \alpha_{m+1})^{-1} \mathbf{T}^0 \cdot \boldsymbol{\alpha} \quad (3.3.2)$$

Se verifica que las visitas sucesivas al estado instantáneo forma un proceso de renovación con distribución $F(\cdot)$, dado por (3.2.1). El proceso puntual obtenido es un proceso de renovación tipo fase.

Definición 3.3.1: La representación $(\boldsymbol{\alpha}, \mathbf{T})$ es denominada irreducible si y sólo si la matriz \mathbf{Q}^* es irreducible.

Lema 3.3.1: Cada componente de $\mathbf{v}(t) = \boldsymbol{\alpha} \exp(\mathbf{T}t)$, $t \geq 0$, es estrictamente positivo para toda $t > 0$, o idénticamente cero para $t \geq 0$, en el último caso la matriz \mathbf{Q}^* es irreducible.

Demostración: Si $T_{ij} \geq 0$, para $i \neq j$, la matriz $\exp(\mathbf{T}t)$ es no negativa para toda $t \geq 0$. Se sigue que $\mathbf{v}(t) \geq \mathbf{0}$, pero entonces $\boldsymbol{\alpha} \neq \mathbf{0}$.

Se supone ahora que en $t_0 > 0$, algunos componentes de $\mathbf{v}(t_0)$ son cero.

Renombramos los índices para que $v_i(t_0) > 0$, para $1 \leq i \leq m_1 < m$, y $v_i(t_0) = 0$, para $m_1 + 1 \leq i \leq m$. Tenemos que

$$v_j'(t_0) = \sum_{i=1}^{m_1} v_i(t_0) T_{ij}, \text{ para } m_1 + 1 \leq j \leq m$$

Si cualquiera de los elementos T_{ij} , $1 \leq i \leq m_1$, $m_1 + 1 \leq j \leq m$, son positivos, entonces para alguna j , $m_1 + 1 \leq j \leq m$, $v_j'(t_0) > 0$, y por lo tanto $v_j(t_0 - \varepsilon) = 0$, $\varepsilon > 0$, lo cual es una contradicción. Por lo tanto la matriz \mathbf{T} es de la forma

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{T}_3 & \mathbf{T}_4 \end{pmatrix}.$$

Ahora se divide \mathbf{a} y $\mathbf{v}(t)$ en $\mathbf{a}_1, \mathbf{a}_2$ y $\mathbf{v}_1(t), \mathbf{v}_2(t)$ respectivamente. Se sigue entonces que

$$\mathbf{v}_2(t) = \mathbf{a}_2 \exp(\mathbf{T}_4 t), \text{ para } t \geq 0.$$

La matriz $\exp(\mathbf{T}_4 t)$ es no singular. Entonces $\mathbf{v}_2(t_0) = \mathbf{0}$, se sigue que $\mathbf{a}_2 = \mathbf{0}$, y por lo tanto $\mathbf{v}_2(t) = \mathbf{0}$, para $t \geq 0$. Esto implica que \mathbf{Q}^* es reducible y

$$F(t) = 1 - \mathbf{a}_1 \exp(\mathbf{T}_1 t) \mathbf{e}_1, \text{ para } t \geq 0. \quad \blacksquare$$

Teorema 3.3.1: Si la matriz \mathbf{Q}^* es reducible, podemos eliminar columnas y renglones de \mathbf{Q} correspondientes al subconjunto de índices de $1, \dots, m$ para obtener una representación más pequeña e irreducible $(\mathbf{a}_1, \mathbf{T}_1)$ de $F(\cdot)$.

Demostración: Si \mathbf{Q}^* es reducible, se puede escribir de la forma

$$\mathbf{Q}^* = \begin{pmatrix} \mathbf{T}_1 + \mathbf{T}_1^0 \mathbf{A}_1^0 & \mathbf{T}_2 + \mathbf{T}_1^0 \mathbf{A}_2^0 \\ \mathbf{T}_3 + \mathbf{T}_2^0 \mathbf{A}_1^0 & \mathbf{T}_4 + \mathbf{T}_2^0 \mathbf{A}_2^0 \end{pmatrix}$$

donde $\mathbf{T}_1 + \mathbf{T}_1^0 \mathbf{A}_1^0$ y $\mathbf{T}_4 + \mathbf{T}_2^0 \mathbf{A}_2^0$ son de orden m_1 y $m - m_1$ respectivamente, y $\mathbf{T}_2 + \mathbf{T}_1^0 \mathbf{A}_2^0 = \mathbf{0}$. Entonces los elementos por debajo de la diagonal de \mathbf{Q}^* son no negativos, se sigue que $\mathbf{T}_2 = \mathbf{0}$, y $\mathbf{T}_1^0 \mathbf{A}_2^0 = \mathbf{0}$.

Si $\alpha_j > 0$, para alguna j con $m_1 + 1 \leq j \leq m$, se sigue que $\mathbf{T}_1^0 = \mathbf{0}$. Entonces \mathbf{T} es de la forma

$$\begin{pmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{T}_3 & \mathbf{T}_4 \end{pmatrix}$$

esto implicaría que la matriz \mathbf{T}_1 y por lo tanto también la matriz \mathbf{T} son singulares. Esto contradice la hipótesis inicial, por lo que se tiene que $\alpha_2 = 0$.

Evaluamos la distribución $F(\cdot)$ y obtenemos

$$F(x) = 1 - (\alpha_1, \mathbf{0}) \exp \left[x \begin{pmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{T}_3 & \mathbf{T}_4 \end{pmatrix} \right] \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix}$$

$$F(t) = 1 - \alpha_1 \exp(\mathbf{T}_1 t) \mathbf{e}_1, \text{ para } t \geq 0.$$

$F(\cdot)$ también tiene la representación más pequeña (α_1, \mathbf{T}_1) . ■

Corolario 3.3.1: Para una representación irreducible (α, \mathbf{T}) , los vectores $\alpha \exp(\mathbf{T}t)$ y $\exp(\mathbf{T}t)\mathbf{T}^0$ son estrictamente positivos para $t > 0$.

Demostración: El mismo argumento muestra que cada componente de ese vector es idénticamente cero o positivo para toda $t > 0$. Si fuese el i -ésimo componente cero, la probabilidad de absorción en el proceso \mathbf{Q} , iniciando en el estado i sería cero. Esto es una contradicción. ■

Corolario 3.3.2: Para $s \geq 0$, los vectores $\alpha(s\mathbf{I} - \mathbf{T})^{-1}$ y $(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{T}^0$ son positivos.

Por lo tanto siempre podemos asumir que una representación (α, \mathbf{T}) de una distribución tipo fase es irreducible.

3.4 Propiedades de Cerradura de las Distribuciones Tipo Fase

Algunas operaciones con las distribuciones tipo fase nos conducen nuevamente a distribuciones tipo fase obteniendo una nueva representación para la distribución tipo fase obtenida. Esto es algorítmicamente útil cuando las operaciones involucradas, las cuales en general requieren integraciones numéricas, pueden ser reemplazadas por operaciones matriciales.

Si \mathbf{T}^0 es un vector columna de tamaño m y β es un vector renglón de tamaño n , denotaremos por $\mathbf{T}^0 \mathbf{B}^0$ a la matriz obtenida del producto escalar $\mathbf{T}^0 \cdot \beta$, con elementos $T_i^0 \cdot \beta_j$.

Probaremos que la familia de las distribuciones tipo fase es cerrada bajo convoluciones y mezclas.

Teorema 3.4.1: Si $F(\cdot)$ y $G(\cdot)$ son distribuciones tipo fase ambas continuas o ambas discretas con representaciones $(\boldsymbol{\alpha}, \mathbf{T})$ y $(\boldsymbol{\beta}, \mathbf{S})$ de órdenes m y n respectivamente, entonces su convolución $F * G(\cdot)$ es una distribución tipo fase con representación $(\boldsymbol{\gamma}, \mathbf{L})$, en el caso continuo dada por:

$$\boldsymbol{\gamma} = [\boldsymbol{\alpha}, \alpha_{m+1} \boldsymbol{\beta}]$$

$$\mathbf{L} = \begin{pmatrix} \mathbf{T} & \mathbf{T}^0 \mathbf{B}^0 \\ \mathbf{0} & \mathbf{S} \end{pmatrix}. \quad (3.4.1)$$

Demostración: La transformada de Laplace de la distribución tipo fase con representación $(\boldsymbol{\gamma}, \mathbf{L})$ está dada por

$$\hat{l}(s) = (\boldsymbol{\alpha}_1)_0 (\boldsymbol{\alpha}_2)_0 + (\boldsymbol{\alpha}_1, (\boldsymbol{\alpha}_1)_0 \boldsymbol{\alpha}_2) \begin{pmatrix} (s\mathbf{I} - \mathbf{T}_1)^{-1} & \mathbf{A} \\ \mathbf{0} & (s\mathbf{I} - \mathbf{T}_2)^{-1} \end{pmatrix} (\tilde{\mathbf{t}}_1, \mathbf{t}_2)$$

$$\hat{l}(s) = \alpha_{m+1} \beta_{n+1} + \beta_{n+1} \boldsymbol{\alpha} (s\mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0 + \alpha_{m+1} \boldsymbol{\beta} (s\mathbf{I} - \mathbf{S})^{-1} \mathbf{S}^0 + \boldsymbol{\alpha} (s\mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0 \mathbf{B}^0 (s\mathbf{I} - \mathbf{S})^{-1} \mathbf{S}^0$$

El último término se simplifica a $\boldsymbol{\alpha} (s\mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0 \cdot \boldsymbol{\beta} (s\mathbf{I} - \mathbf{S})^{-1} \mathbf{S}^0$, por lo tanto la transformada de Laplace es igual a

$$\hat{l}(s) = [\alpha_{m+1} + \boldsymbol{\alpha} (s\mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0] [\beta_{n+1} + \boldsymbol{\beta} (s\mathbf{I} - \mathbf{S})^{-1} \mathbf{S}^0]$$

$$\hat{l}(s) = \hat{l}_1(s) \hat{l}_2(s).$$

■

Es interesante notar la relación del proceso de Markov con matriz

$$\begin{pmatrix} \mathbf{T} & \mathbf{T}^0 \mathbf{B}^0 \\ \mathbf{S}^0 \mathbf{A}^0 & \mathbf{S} \end{pmatrix}$$

con el proceso de renovación alternante con distribuciones $F(\cdot)$ y $G(\cdot)$.

Si al tiempo t el proceso de Markov está en el conjunto de estados $\{1, \dots, m\}$, el punto t es cubierto por un intervalo con distribución $F(\cdot)$. Una interpretación similar se obtiene para la estancia en el conjunto $\{m+1, \dots, m+n\}$. Transiciones del conjunto $\{1, \dots, m\}$ al conjunto $\{m+1, \dots, m+n\}$, y viceversa, corresponden a las renovaciones.

Definimos el vector $\boldsymbol{\pi}$ como el vector de probabilidad estacionario del generador irreducible (3.3.1). Claramente el vector positivo $\boldsymbol{\pi}$ es la única solución a las ecuaciones

$$\boldsymbol{\pi}(\mathbf{T} + \mathbf{T}^0 \mathbf{A}^0) = \mathbf{0}, \quad \boldsymbol{\pi} \mathbf{e} = 1 \quad (3.4.3)$$

Las siguientes fórmulas serán utilizadas posteriormente

$$\boldsymbol{\pi} = -\boldsymbol{\pi}\mathbf{T}^0\mathbf{A}^0\mathbf{T}^{-1} = -(1-\alpha_{m+1})^{-1}(\boldsymbol{\pi}\mathbf{T}^0)\boldsymbol{\alpha}\mathbf{T}^{-1}, \quad (3.4.4)$$

$$\boldsymbol{\pi}\mathbf{T}^0 = (1-\alpha_{m+1})^{-1}\mu_1^{\prime-1} \quad (3.4.5)$$

donde μ_1^{\prime} es la esperanza de $F(\cdot)$.

Teorema 3.4.2: Si $F(\cdot)$ es una distribución tipo fase con representación $(\boldsymbol{\alpha}, \mathbf{T})$, entonces

$$F^*(x) = \frac{1}{\mu_1^{\prime}} \int_0^x [1 - F(u)] du, \text{ para } x \geq 0$$

es una distribución tipo fase con representación $(\boldsymbol{\pi}, \mathbf{T})$.

Demostración:

$$F^*(x) = \frac{1}{\mu_1^{\prime}} \int_0^x \boldsymbol{\alpha} \exp(\mathbf{T}u) du \mathbf{e} = \frac{1}{\mu_1^{\prime}} \boldsymbol{\alpha} \mathbf{T}^{-1} [\exp(\mathbf{T}x) - \mathbf{I}] \mathbf{e}$$

$$F^*(x) = \mathbf{1} - \boldsymbol{\pi} \exp(\mathbf{T}x) \mathbf{e},$$

por la función generadora de momentos $\mu_i^{\prime} = (-1)^n n! \boldsymbol{\alpha} \mathbf{T}^{-n} \mathbf{e}$ y por (3.4.4). ■

Los momentos μ_i^* de $F^*(\cdot)$ están dados por

$$\mu_i^* = (-1)^i i! \boldsymbol{\pi} \mathbf{T}^{-1} \mathbf{e} = (-1)^{i+1} \frac{i!}{\mu_1^{\prime}} \boldsymbol{\alpha} \mathbf{T}^{-i-1} \mathbf{e}$$

$$\mu_i^* = \frac{\mu_{i+1}^{\prime}}{(i+1)\mu_1^{\prime}}, \text{ para } i \geq 1.$$

Teorema 3.4.3: La convolución de dos distribuciones tipo fase $PH(\boldsymbol{\alpha}_1, \mathbf{T}_1, E_1)$ y $PH(\boldsymbol{\alpha}_2, \mathbf{T}_2, E_2)$ es una distribución tipo fase con parámetros $(\boldsymbol{\alpha}, \mathbf{T}, E)$, donde $E = E_1 \cup E_2$,

$$\alpha_i = \begin{cases} (\boldsymbol{\alpha}_1)_i & \text{si } i \in E_1 \\ (\boldsymbol{\alpha}_1)_0 (\boldsymbol{\alpha}_2)_i & \text{si } i \in E_2 \end{cases}$$

y

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_1 & \mathbf{t}_1 \boldsymbol{\alpha}_2 \\ \mathbf{0} & \mathbf{T}_2 \end{pmatrix}$$

donde $(\mathbf{a}_k)_i$ es el i -ésimo componente de \mathbf{a}_k , y $\mathbf{t}_k = -\mathbf{T}_k \mathbf{e}$, $k = 1, 2$.

Demostración: Sean $\hat{l}(s)$, $\hat{l}_1(s)$, $\hat{l}_2(s)$, las transformadas de Laplace de $PH(\mathbf{a}, \mathbf{B}, E)$, $PH(\mathbf{a}_1, \mathbf{T}_1, E_1)$ y $PH(\mathbf{a}_2, \mathbf{T}_2, E_2)$ respectivamente. Entonces, es suficiente probar que $\hat{l}(s) = \hat{l}_1(s)\hat{l}_2(s)$ para $s \geq 0$. Notemos que

$$s\mathbf{I} - \mathbf{T} = \begin{pmatrix} s\mathbf{I} - \mathbf{T}_1 & -\mathbf{t}_1\mathbf{a}_2 \\ \mathbf{0} & s\mathbf{I} - \mathbf{T}_2 \end{pmatrix}.$$

Primero demostraremos que la matriz $s\mathbf{I} - \mathbf{T}$ es invertible. Esto es equivalente a mostrar que existe una matriz \mathbf{A} para la cual

$$\begin{pmatrix} s\mathbf{I} - \mathbf{T}_1 & -\mathbf{t}_1\mathbf{a}_2 \\ \mathbf{0} & s\mathbf{I} - \mathbf{T}_2 \end{pmatrix} \times \begin{pmatrix} (s\mathbf{I} - \mathbf{T}_1)^{-1} & \mathbf{A} \\ \mathbf{0} & (s\mathbf{I} - \mathbf{T}_2)^{-1} \end{pmatrix} = \mathbf{I}.$$

Es decir, \mathbf{A} debe satisfacer $(s\mathbf{I} - \mathbf{T}_1)\mathbf{A} - \mathbf{t}_1\mathbf{a}_2(s\mathbf{I} - \mathbf{T}_2)^{-1} = \mathbf{0}$

$$(s\mathbf{I} - \mathbf{T}_1)^{-1} = \begin{pmatrix} (s\mathbf{I} - \mathbf{T}_1)^{-1} & \mathbf{A} \\ \mathbf{0} & (s\mathbf{I} - \mathbf{T}_2)^{-1} \end{pmatrix},$$

donde $\mathbf{A} = (s\mathbf{I} - \mathbf{T}_1)^{-1}\mathbf{t}_1\mathbf{a}_2(s\mathbf{I} - \mathbf{T}_2)^{-1}$. Sea $\tilde{\mathbf{t}}_1$ el vector que satisface $\tilde{\mathbf{t}}_1 + (\mathbf{T}_1 + \mathbf{a}_2)\mathbf{e} = \mathbf{0}$ y es equivalente $\tilde{\mathbf{t}}_1 = (\mathbf{a}_2)_0\mathbf{t}_1$

Utilizando la transformada de Laplace $\hat{l}(s)$ de $PH(\mathbf{a}, \mathbf{T}, E)$

$$\hat{l}(s) = (\mathbf{a}_1)_0(\mathbf{a}_2)_0 + (\mathbf{a}_1, (\mathbf{a}_1)_0\mathbf{a}_2) \begin{pmatrix} (s\mathbf{I} - \mathbf{T}_1)^{-1} & \mathbf{A} \\ \mathbf{0} & (s\mathbf{I} - \mathbf{T}_2)^{-1} \end{pmatrix} (\tilde{\mathbf{t}}_1, \mathbf{t}_2)$$

$$\hat{l}(s) = (\mathbf{a}_1)_0(\mathbf{a}_2)_0 + \mathbf{a}_1(s\mathbf{I} - \mathbf{T}_1)^{-1}\tilde{\mathbf{t}}_1 + \mathbf{a}_1\mathbf{A}\mathbf{t}_2 + (\mathbf{a}_1)_0(\mathbf{a}_2)(s\mathbf{I} - \mathbf{T}_2)^{-1}\mathbf{t}_2$$

$$\hat{l}(s) = ((\mathbf{a}_1)_0 + \mathbf{a}_1(s\mathbf{I} - \mathbf{T}_1)^{-1}\tilde{\mathbf{t}}_1)((\mathbf{a}_2)_0 + \mathbf{a}_2(s\mathbf{I} - \mathbf{T}_2)^{-1}\mathbf{t}_2)$$

$$\hat{l}(s) = \hat{l}_1(s)\hat{l}_2(s).$$

■

Este teorema puede ser generalizado para la convolución $PH(\mathbf{a}_1, \mathbf{T}_1, E_1) * \dots * PH(\mathbf{a}_n, \mathbf{T}_n, E_n)$ de n distribuciones tipo fase es una distribución tipo fase para $n \geq 2$. Sus parámetros $(\mathbf{a}, \mathbf{T}, E)$

están dados por $E = \bigcup_{k=1}^n E_k$, $\prod_{j=1}^{k-1} (\mathbf{a}_j)_0 (\mathbf{a}_k)_i$ si $i \in E_k$, donde $k = 1, \dots, n$.

Las mezclas infinitas de distribuciones tipo fase en general no son de tipo fase.

Teorema 3.4.4: Sea $0 < p < 1$. La mezcla $pPH(\boldsymbol{\alpha}_1, \mathbf{T}_1, E_1) + (1-p)PH(\boldsymbol{\alpha}_2, \mathbf{T}_2, E_2)$ es una distribución tipo fase con parámetros $(\boldsymbol{\alpha}, \mathbf{T}, E)$ donde $E = E_1 \cup E_2$

$$\alpha_i = \begin{cases} p(\alpha_1)_i & \text{si } i \in E_1 \\ (1-p)(\alpha_2)_i & \text{si } i \in E_2 \end{cases}$$

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 \end{pmatrix}.$$

Este teorema implica que para $n \geq 2$ y para cualquier función de probabilidad (p_1, \dots, p_n) , la mezcla $\sum_{k=1}^n p_k PH(\boldsymbol{\alpha}_k, \mathbf{T}_k, E_k)$ es una distribución fase con parámetros $(\boldsymbol{\alpha}, \mathbf{T}, E)$, donde

$$E = \bigcup_{k=1}^n E_k, \quad \alpha_i = p_k (\alpha_k)_i \text{ si } i \in E_k \text{ y}$$

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{T}_n \end{pmatrix}.$$

Mostraremos que la clase de las distribuciones tipo fase es densa en la clase de todas las distribuciones de variables aleatorias no negativas. Es decir, para cualquier distribución F definida en \mathbb{R}_+ , existe una sucesión de distribuciones tipo fase F_n tales que $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ para cada punto de continuidad x de F .

Lema 3.4.1: Sea $\{X_n\}$ una sucesión de variables aleatorias evaluadas en \mathbb{R} y $x \in \mathbb{R}$. Si $\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = x$ y $\lim_{n \rightarrow \infty} \mathbb{E}(X_n^2) = x^2$, entonces

$$\lim_{n \rightarrow \infty} \mathbb{E}(f(X_n)) = f(x)$$

para cada función acotada $f: \mathbb{R} \rightarrow \mathbb{R}$ siendo continua en x .

Demostración: Sin pérdida de generalidad, suponemos que $\sup_y |f(y)| \leq 1$. Para cada $\varepsilon > 0$ elegimos $\delta > 0$ tal que $|f(x) - f(y)| \leq \varepsilon$ siempre que $|x - y| \leq \delta$. Entonces

$$|\mathbb{E}(f(X_n) - f(x))| \leq \mathbb{E}|f(X_n) - f(x)|$$

$$\mathbb{E}|f(X_n) - f(x)| \leq \mathbb{E}[|f(X_n) - f(X)|; |X_n - x| \leq \delta] + \mathbb{E}[|f(X_n) - f(X)|; |X_n - x| > \delta]$$

$$\mathbb{E}|f(X_n) - f(x)| \leq \varepsilon + 2\mathbb{P}(|X_n - x| > \delta)$$

Por la desigualdad de Chebyshev

$$\mathbb{P}(|X_n - x| > \delta) \leq \frac{\mathbb{E}(X_n - x)^2}{\delta^2} = \frac{\mathbb{E}(X_n^2) - x^2 + 2x(x - \mathbb{E}(X_n))}{\delta^2}.$$

■

Teorema 3.4.5: La familia de las distribuciones Tipo Fase es densa en el conjunto de todas las distribuciones en \mathbb{R}_+ .

Demostración: Sea F una distribución arbitraria en \mathbb{R}_+ , definimos

$$F_n = F(0)\delta_0 + \sum_{k=1}^{\infty} \left(F\left(\frac{k}{n}\right) - F\left(\frac{k-1}{n}\right) \right) \text{Erl}(k, n)$$

para $n \geq 1$. Mostraremos que $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ para cada $x \geq 0$ con $F(x) = F(x-)$. Notemos que

$$F_n(x) = \sum_{k=0}^{\infty} F\left(\frac{k}{n}\right) e^{-nx} \frac{(nx)^k}{k!} = \int_0^{\infty} F(t) dG_{n,x}(t)$$

$$\text{donde } G_{n,x} = \sum_{k=0}^{\infty} e^{-nx} \frac{(nx)^k}{k!} \delta_{\frac{k}{n}}.$$

Además

$$\int_0^{\infty} t dG_{n,x}(t) = \sum_{k=0}^{\infty} \frac{k}{n} \frac{(nx)^k}{k!} e^{-nx} = x \sum_{k=1}^{\infty} \frac{(nx)^{k-1}}{(k-1)!} e^{-nx} = x$$

y

$$\int_0^{\infty} t^2 dG_{n,x}(t) = n^{-1}x + x^2 \rightarrow x^2 \text{ conforme } n \rightarrow \infty.$$

Entonces por el Lema 3.4.1

$$\int_0^{\infty} F(t) dG_{n,x}(t) = E(F(X_n)) \rightarrow F(X),$$

donde X_n tiene distribución $G_{n,x}$. Ahora sea

$$F'_n = (F(0) + \bar{F}(n)) \delta_0 + \sum_{k=1}^{n^2} \left(F\left(\frac{k}{n}\right) - F\left(\frac{k-1}{n}\right) \right) \text{Erl}(k, n).$$

Entonces cada distribución Erlang es tipo fase. ■

3.5 Reclamaciones Tipo Fase

Consideremos un proceso de riesgo (Bladt 2004).

$$R_t = u + pt - \sum_{n=0}^{N_t} U_n$$

donde N_t es un proceso Poisson con intensidad β , las reclamaciones U_1, U_2, \dots son variables aleatorias independientes idénticamente distribuidas tipo fase $(\boldsymbol{\pi}, \mathbf{T})$ de dimensión d . Suponemos sin pérdida de generalidad que $p = 1$.

Consideramos el proceso de superávit de las reclamaciones como

$$S_t = u - R_t = \sum_{n=0}^{N_t} U_n - t.$$

La probabilidad de ruina se puede escribir como

$$\psi(u) = \mathbb{P}\left(\sup_{t \geq 0} S_t > u\right).$$

Sea $\tau_+ = \inf(t \geq 0 | S_t > 0)$. Cuando $S_t > 0$ por primera vez, ocurre una reclamación. Esta reclamación está generada por un proceso de Markov subyacente $\{M_t\}$. Iniciando el proceso desde el nivel $S_{(\tau_+)}$ y corriendo en sentido vertical, el proceso subyacente cruzara el nivel cero al tiempo $S_{(\tau_+)}$. Al cruzar el nivel cero el proceso subyacente $\{M_t\}$ se encuentra en uno de los estados $1, 2, \dots, d$. Definimos $v_i = \mathbb{P}(M_{S_{(\tau_+)}} = i)$ i.e. la probabilidad que al cruzar el nivel 0 el proceso $\{M_t\}$ lo hará en el estado i .

Definimos

$$\tau_+(n+1) = \inf \{t \geq 0 \mid S_t > S_{\tau_+(n)}\}$$

$\tau_+(n)$ es la n -ésima vez que S_t llegará a un nuevo nivel máximo global.

La distribución del exceso $S_{\tau_+(n+1)} - S_{\tau_+(n)}$ es tipo fase (\mathbf{v}, \mathbf{T}) y el proceso escalonado ascendente forma un proceso de renovación propio y defectuoso con distribución entre llegadas tipo fase (\mathbf{v}, \mathbf{T}) . La ruina ocurre si y solo si la vida del proceso de renovación rebasara el nivel u lo cual es equivalente a que el proceso de renovación se encuentre en uno de los estados $1, 2, \dots, d$ al tiempo u . Esta probabilidad es

$$\mathbf{v} \exp((\mathbf{T} + \mathbf{t}\mathbf{v})u) \mathbf{e},$$

la cual es la probabilidad de ruina.

Lema 3.5.1: Para cualquier $t > 0$

$$S_u^* = S_t - S_{t-u} \stackrel{D}{=} S_u$$

Demostración:

$$S_u^* = S_t - S_{t-u}$$

$$S_u^* = \sum_{j=0}^{N(t)} Y_j - pt - \left(\sum_{j=0}^{N(t-u)} Y_j - p(t-u) \right)$$

$$S_u^* \stackrel{D}{=} \sum_{j=0}^{N(u)} Y_j - pu$$

$$S_u^* = S_u$$

Lema 3.5.2: $R_+(\cdot)$ es $\frac{1}{p}$ veces la medida de Lebesgue restringido a $(-\infty, 0)$

Demostración: Notemos que

$$R_+(A) = \mathbb{E} \int_0^\infty I\{S_t \in A, \tau_+ > t\} dt = \int_0^\infty \mathbb{P}(S_t \in A, \tau_+ > t) dt$$

Por Lema (3.5.1) y usando que $S_u = S_t^* - S_{t-u}^*$ tenemos

$$\mathbb{P}(S_t \in A, \tau_+ > t) = \mathbb{P}(S_t \in A, S_u \leq 0, 0 \leq u \leq t)$$

$$\mathbb{P}(S_t \in A, \tau_+ > t) = \mathbb{P}(S_t^* \in A, S_t^* - S_{t-u}^* \leq 0, 0 \leq u \leq t)$$

$$\mathbb{P}(S_t \in A, \tau_+ > t) = \mathbb{P}(S_t^* \in A, S_t^* \leq S_{t-u}^* \leq 0, 0 \leq u \leq t)$$

$$\mathbb{P}(S_t \in A, \tau_+ > t) = \mathbb{P}(S_t \in A, S_t \leq S_{t-u} \leq 0, 0 \leq u \leq t)$$

$$\mathbb{P}(S_t \in A, \tau_+ > t) = \mathbb{P}(S_t \in A, S_t \leq S_u \leq 0, 0 \leq u \leq t)$$

Entonces

$$R_+(A) = \int_0^\infty \mathbb{P}(S_t \in A, S_t \leq S_{t-u} \leq 0, 0 \leq u \leq t) dt$$

$$R_+(A) = \mathbb{E} \int_0^\infty I\{S_t \in A, S_t \leq S_{t-u} \leq 0, 0 \leq u \leq t\} dt$$

Entonces $R_+(A)$ es el tiempo esperado en el que S_t se encuentra en A y en un valor mínimo. Como $S_t \rightarrow -\infty$ y S_t es continua fuera de los saltos, entonces eventualmente S_t pasará por todos los puntos en A . La pendiente de S_t fuera de los saltos es de $-p$ y pasará por A p veces más rápido. ■

Teorema 3.5.1: La probabilidad de ruina está dada por

$$\psi(u) = \pi_+ \exp((\mathbf{T} + \mathbf{t}\pi)u)\mathbf{e} \text{ donde } \pi_+ = -\beta\pi\mathbf{T}^{-1}.$$

Demostración: Un proceso subyacente cruzará el nivel cero en el estado j por primera vez a tiempo t si y solo si $t \leq \tau_u$ y si el proceso subyacente pasara de algún estado inicial i al estado j durante el tiempo $-S_{t-}$. La probabilidad de que en $[t, t+dt)$ habrá una llegada es de βdt . Entonces por la ley de la probabilidad total

$$v_i = \int_0^\infty \mathbb{E} \left[\sum_{i=1}^d \pi_i p_{ij}^{-S_{t-}} I\{\tau_+ \leq t\} \right] \beta dt$$

$$v_i = \beta \sum_{i=1}^d \pi_i \int_0^\infty \mathbb{E} \left[p_{ij}^{-S_{t-}} I\{\tau_+ \leq t\} \right] dt$$

$$v_i = \beta \sum_{i=1}^d \pi_i \mathbb{E} \left[\int_0^{\tau_+} p_{ij}^{-S_{t-}} dt \right]$$

Definimos

$$R_+(A) = \mathbb{E} \left[\int_0^{\tau_+} \mathbb{1}_{S_t \in A} dt \right]$$

Notemos que R_+ está concentrado sobre $\mathbb{R}_- = (-\infty, 0)$. En teoría de la medida para cualquier función medible no negativa tenemos que

$$\int_{-\infty}^0 f(y) R_+(dy) = \mathbb{E} \left[\int_0^{\tau_+} f(S_t) dt \right].$$

Entonces

$$v_i = \beta \sum_{i=1}^d \pi_i \mathbb{E} \left[\int_0^{\tau_+} p_{ij}^{-S_t} dt \right]$$

$$v_i = \beta \sum_{i=1}^d \pi_i \int_{-\infty}^0 p_{ij}^{-y} R_+(dy)$$

$$v_i = \beta \sum_{i=1}^d \pi_i \int_0^{\infty} p_{ij}^y R_+(-dy)$$

$$v_i = \int_0^{\infty} \left(\beta \sum_{i=1}^d \pi_i p_{ij}^y \right) R_+(-dy)$$

donde la primera ecuación resulta por la probabilidad de un arribo a tiempo t es cero. Por el Lema 3.5.2, $R_+(-dy) = \lambda(dy)$ donde λ es la medida de Lebesgue restringido a $(-\infty, 0)$. Entonces

$$v_i = \int_0^{\infty} \beta \sum_{i=1}^d \pi_i p_{ij}^y R_+(-dy)$$

$$v_i = \int_0^{\infty} \beta \sum_{i=1}^d \pi_i p_{ij}^y dy$$

resultando en

$$\mathbf{v} = \int_0^{\infty} \beta \boldsymbol{\pi} e^{\mathbf{T}y} dy = -\beta \boldsymbol{\pi} \mathbf{T}^{-1} \quad \blacksquare$$

3.6 Ejemplos de Distribuciones Tipo Fase

La distribución exponencial con función de densidad $f(u) = \lambda e^{-\lambda u}$ tiene una representación

$$\boldsymbol{\alpha} = (1)$$

$$\mathbf{T} = (-\lambda)$$

La distribución hiperexponencial con función de densidad $f(u) = \sum_{i=1}^p \alpha_i \lambda_i e^{-\lambda_i u}$ donde, para $i = 1, \dots, p$, $\alpha_i > 0$ y $\sum_{i=1}^p \alpha_i = 1$, tiene una representación

$$\mathbf{\alpha} = (\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_p)$$

$$\mathbf{T} = \begin{pmatrix} -\lambda_1 & 0 & \dots & 0 \\ 0 & -\lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -\lambda_p \end{pmatrix}$$

La distribución Erlang con parámetro p con función de densidad $f(u) = \frac{\lambda^p u^{p-1} e^{-\lambda u}}{p!}$ tiene una representación

$$\mathbf{\alpha} = (1 \quad 0 \quad \dots \quad 0)$$

$$\mathbf{T} = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & 0 \\ 0 & -\lambda & \lambda & \dots & 0 \\ 0 & 0 & -\lambda & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda \end{pmatrix}$$

La distribución Coxian con parámetro p tiene representaciones de la forma

$$\mathbf{\alpha} = (\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_p)$$

$$\mathbf{T} = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & \dots & 0 \\ 0 & -\lambda_2 & \lambda_2 & \dots & 0 \\ 0 & 0 & -\lambda_3 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_p \end{pmatrix}$$

donde $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_p$

3.7 Distribuciones Tipo Fase Multivariadas

Existen diversas definiciones concernientes a distribuciones multivariadas de tipo exponencial o gamma. Estas distribuciones tienen distribuciones marginales exponenciales o gamma. Un número de estas distribuciones marginales tienen una transformada de Laplace racional multidimensional. También, la clase de las distribuciones tipo fase, la cual es la generalización de cierto tipo de distribuciones gamma, han sido extendidas al ámbito multivariado por Assaf. et al. (1984) y posteriormente por Kulkarni (1989). La última clase, la cual contiene un caso especial, provee una construcción elegante de distribuciones tipo fase multivariadas en términos de un proceso subyacente de Markov de salto. Además tiene una generalización natural a las distribuciones matriz exponenciales multivariadas.

Assaf et al. (1984) introdujo una clase de distribuciones tipo fase multivariadas, considerando los tiempos de llegada a diferentes subconjuntos del espacio de estados. Más específicamente, consideramos un generador tipo fase (matriz de subintensidades) \mathbf{T} de dimensión m y sea $\{J_t\}_{t \geq 0}$ el proceso de Markov subyacente de salto. Sea Γ_i , $i=1,2,\dots,n$ los subconjuntos absorbentes del espacio de estados. Sea X_i el primer tiempo de llegada de $\{J_t\}_{t \geq 0}$ a Γ_i . Entonces el vector n -dimensional $\mathbf{X} = (X_1, \dots, X_n)$ tiene una distribución tipo fase en la clase de las distribuciones tipo fase multivariadas.

El tiempo de estancia para X_k es acumulada con tasa de 1 en estados pertenecientes a Γ_k^c , donde Γ_k^c es el complemento de Γ_k . Basado en esta interpretación, Kulkarni (1989) introdujo una generalización de la clase de las distribuciones tipo fase multivariadas. En esta clase el tiempo de estancia X_i es acumulado en el estado j con tasa \mathbf{K}_{ij} . No existe alguna restricción para el generador tipo fase \mathbf{T} . Si el tiempo total de estancia en el estado j antes de la absorción es denotado por \mathbf{Y}_j , definimos un vector aleatorio n -dimensional $\mathbf{X}_i = \sum_{j=1}^m \mathbf{K}_{ij} \mathbf{Y}_j$.

El siguiente teorema nos proporciona una caracterización alternativa en términos de todas las proyecciones no negativas.

Teorema 3.7.1: Una distribución en la generalización de la clase de las distribuciones tipo fase multivariadas puede ser caracterizada por $\langle \mathbf{a}, \mathbf{X} \rangle$ estando distribuidas tipo fase con representación $(\mathbf{a}, \mathbf{T}(\mathbf{a}))$ con $\mathbf{T}(\mathbf{a}) = \mathbf{\Lambda}(\mathbf{K}\mathbf{a})^{-1}\mathbf{T}$, donde $\mathbf{\Lambda}(b)$ es la matriz diagonal con b en la diagonal.

Demostración: Sea \mathbf{X} una variable aleatoria con una distribución la generalización de la clase de las distribuciones tipo fase multivariadas. Entonces el i -ésimo componente X_i de \mathbf{X} puede ser escrito como

$$\mathbf{X}_i = \sum_{j=1}^m \sum_{k=1}^{N_k} \mathbf{K}_{ij} \mathbf{Z}_{jk}$$

donde N_k denota el número de visitas al estado transitorio k en la cadena de Markov a tiempo continuo $m+1$ -dimensional con m estados transitorios y un estado absorbente. Así, esta cadena de Markov define una distribución tipo fase a tiempo continuo. Denotamos la parte transitoria de la matriz generadora por \mathbf{T} . Las variables aleatorias \mathbf{Z}_{jk} son la k -ésima estancia en el estado j ,

mientras \mathbf{K}_{ij} son constantes reales no negativas. Considerando las distribuciones de la familia de proyecciones dadas por $\langle \mathbf{a}, \mathbf{X} \rangle$ obtenemos

$$\langle \mathbf{a}, \mathbf{X} \rangle = \sum_{i=1}^n a_i \sum_{j=1}^m \sum_{k=1}^{N_k} K_{ij} Z_{jk} = \sum_{j=1}^m \left(\sum_{i=1}^n K_{ij} a_i \right) Z_j$$

con $Z_j = \sum_{k=1}^{N_k} Z_{jk}$. De aquí en adelante asumiremos que $\mathbf{Ka} > 0$. Si esta condición no se satisface podríamos obtener una representación tipo fase inadecuada. Bajo la condición $\mathbf{Ka} > 0$ observamos que $\langle \mathbf{a}, \mathbf{X} \rangle$ se distribuye tipo fase con matriz generadora $\Delta(\mathbf{Ka})^{-1} \mathbf{T}$. ■

En este capítulo se mostraron las distribuciones tipo fase univariadas y multivariadas, sus propiedades, representaciones y caracterizaciones; así como también se describió un proceso de reclamaciones tipo fase. A lo largo de este capítulo se observó que las distribuciones tipo fase son una clase versátil de distribuciones definidas en los números reales no negativos que agregan flexibilidad al modelado estocástico en distintas áreas.

Capítulo 4: Distribuciones Matriz Exponenciales

4.1 Introducción

Las distribuciones matriz exponenciales son distribuciones definidas en los números reales no negativos y con transformada de Laplace racional. La clase de las distribuciones tipo fase es un subconjunto propio de las distribuciones matriz exponenciales. Muchos resultados válidos para las distribuciones tipo fase también son válidos para las distribuciones matriz exponenciales.

Las distribuciones matriz exponenciales tienen funciones de distribución de la misma forma que las distribuciones tipo fase pero sus representaciones no necesariamente tienen interpretaciones probabilísticas simples.

Dada una función racional o una forma equivalente matriz exponencial, no hay procedimientos factibles para asegurar cuáles de ellos corresponden a distribuciones de probabilidad.

Por otro lado, la representación de una distribución tipo fase puede requerir de un número excesivo de parámetros, mientras que las representaciones matriz exponenciales pueden estar dadas por distribuciones de probabilidad con transformadas de Laplace racionales.

También la dimensión es una cuestión importante en las aplicaciones. Generalmente las representaciones matriz exponenciales de distribuciones tipo fase serán de orden menor que sus correspondientes representaciones tipo fase. Por lo tanto es benéfico representar distribuciones tipo fase con matriz exponenciales de orden menor.

4.2 Definición y Propiedades Básicas de las Distribuciones Matriz Exponenciales

Una variable aleatoria X se distribuye conforme a una distribución matriz exponencial si la función de distribución, definida para $x \geq 0$ de la forma (Asmussen y Bladt. 1997):

$$B(x) = \begin{cases} \alpha_0, & x = 0 \\ 1 + \boldsymbol{\alpha} \exp(\mathbf{T}x) \mathbf{T}^{-1} \mathbf{s}, & x > 0 \end{cases} \quad (4.2.1)$$

donde, para $n \geq 1$, $\boldsymbol{\alpha}$ es un vector renglón de $1 \times n$, \mathbf{T} es una matriz de $n \times n$, y \mathbf{s} es un vector columna de $n \times 1$. Es claro $0 \leq \alpha_0 \leq 1$. Además α_0 es la función de distribución continua por la derecha para $x = 0$. Esto es

$$\lim_{x \rightarrow 0^+} (1 + \boldsymbol{\alpha} \exp(\mathbf{T}x) \mathbf{T}^{-1} \mathbf{s}) = \alpha_0.$$

La transformada de Laplace de $B(x)$, definida para $s \in \mathbb{C}$, está dada por

$$\hat{f}(s) = \int_0^{\infty} e^{-sx} dB(x) = \boldsymbol{\alpha} (s\mathbf{I} - \mathbf{T})^{-1} \mathbf{s} + \alpha_0$$

Proposición 4.2.1: Las siguientes proposiciones son equivalentes

$$\hat{B}[s] \text{ es una función racional} \quad (4.2.2)$$

$$b(x) = \alpha e^{\mathbf{T}x} \mathbf{s}; \quad (4.2.3)$$

$$b(x) = \sum_{j=1}^q c_j x^j e^{\eta_j x}; \quad (4.2.4)$$

$$b(x) = \sum_{j=1}^{q_1} c_j x^j e^{\eta_j x} + \sum_{j=1}^{q_2} d_j x^j \cos(a_j x) e^{\delta_j x} + \sum_{j=1}^{q_3} e_j x^j \sin(b_j x) e^{\zeta_j x}; \quad (4.2.5)$$

donde los parámetros en (4.2.3) y (4.2.4) pueden ser valores complejos mientras que los parámetros en (4.2.5) son valores reales.

Demostración: La demostración de que (4.2.2) y (4.2.3) son equivalentes se sigue de la Proposición (4.2.3) (Ver a continuación). Si $\hat{B}[s]$ es racional, se sigue por expansión fraccional y transformada inversa que (4.2.3) ocurre; la inversa se sigue haciendo el cálculo explícito

$$\int_0^{\infty} e^{-sx} x^j e^{\eta x} dx = (-1) \frac{j!}{(\eta + s)^{j+1}},$$

por la fórmula de Euler

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}, \quad \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}. \quad (4.2.6)$$

Por otra parte (4.2.4) implica (4.2.5), mientras que la inversa es trivial. ■

Usaremos (4.2.3) como la caracterización básica, nos referiremos al vector $(\alpha, \mathbf{T}, \mathbf{s})$ como la representación de la distribución B para una distribución matriz exponencial en K y p para la dimensión de la representación.

Algunas condiciones necesarias para saber cuándo una representación $(\alpha, \mathbf{T}, \mathbf{s})$ define una distribución matriz exponencial son dadas fácilmente, por ejemplo $\alpha \mathbf{T}^{-1} \mathbf{s} = 1$ y $\alpha \mathbf{s} \geq 0$.

Es bien conocido que una distribución tipo fase es matriz exponencial con representación $(\alpha, \mathbf{T}, \mathbf{s})$ donde $\mathbf{s} = -\mathbf{T} \mathbf{e}$ es el vector de tasa de salida.

Las siguientes propiedades de las distribuciones matriz exponenciales, las cuales son análogas de las fórmulas dadas para las distribuciones tipo fase.

Proposición 4.2.2: Sea $B \in K$ con representación $(\alpha, \mathbf{T}, \mathbf{s})$. Entonces:

(a) La función de distribución es $B(x) = 1 + \alpha e^{\mathbf{T}x} \mathbf{T}^{-1} \mathbf{s}$

(b) La función generadora de momentos está dada por $\hat{B}[s] = \alpha (-s \mathbf{I} - \mathbf{T})^{-1} \mathbf{s}$

(c) El n -ésimo momento está dado por $(-1)^{n+1} n! \alpha \mathbf{T}^{-n-1} \mathbf{s}$

Demostración:

(a) se sigue de la integración de la densidad y la fórmula $\int e^{\mathbf{T}x} dx = e^{\mathbf{T}x} \mathbf{T}^{-1} = \mathbf{T}^{-1} e^{\mathbf{T}x}$; mientras que
 (b) se sigue de

$$\hat{B}[s] = \int_0^\infty e^{sx} \boldsymbol{\alpha} e^{\mathbf{T}x} dx = \boldsymbol{\alpha} \int_0^\infty e^{(s\mathbf{I} + \mathbf{T})x} dx$$

$$\hat{B}[s] = \boldsymbol{\alpha} \int_0^\infty e^{(s\mathbf{I} + \mathbf{T})x} dx$$

$$\hat{B}[s] = \boldsymbol{\alpha} (-s\mathbf{I} - \mathbf{T})^{-1} \mathbf{s}.$$

Para (c), se diferencia la función generadora de momentos n veces, usando (b), y sustituyendo $s = 0$. ■

La representación de una distribución matriz exponencial no es única. Una representación $(\boldsymbol{\alpha}, \mathbf{T}, \mathbf{s})$ de una distribución matriz exponencial B es llamada minimal si tiene la menor dimensión posible.

Proposición 4.2.3: La transformada de Laplace de una distribución matriz exponencial puede ser escrita como

$$\hat{f}(s) = \frac{b_1 + b_2s + b_3s^2 + \dots + b_n s^{n-1}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n}, \tag{4.2.7}$$

para alguna $n \geq 1$ y para algunas constantes $a_1, \dots, a_n, b_1, \dots, b_n$. Notando que $a_n = b_1$. Nos referimos a los polinomios $b(s) = b_1 + b_2s + b_3s^2 + \dots + b_n s^{n-1}$ como el numerador y a $a(s) = s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n$ como el denominador de la transformada de Laplace. Definimos las raíces de la transformada de Laplace como las raíces de $b(s)$ y los polos de la transformada de Laplace como las raíces de $a(s)$. Además, la función de densidad tiene la siguiente representación

$$f(s) = \boldsymbol{\alpha} e^{\mathbf{T}x} \mathbf{t}$$

donde

$$\boldsymbol{\alpha} = (b_1, b_2, \dots, b_n)$$

$$\mathbf{T} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{pmatrix}$$

$$\mathbf{t} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

■

Esta proposición es muy importante debido a que es el primer vínculo de la expresión de la transformada de Laplace y por lo tanto también la función generadora de momentos en una representación; notemos también que la ecuación característica de una matriz de la forma

$$\mathbf{T} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix}$$

es

$$s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n = 0$$

los valores propios de \mathbf{T} son idénticos a las raíces de la transformada de Laplace, además muestra que siempre es posible encontrar una representación sencilla de una distribución matriz exponencial; por último la razón más importante es, que existe una correspondencia uno a uno entre la transformada de Laplace de una distribución matriz exponencial y una representación minimal de la forma

$$\boldsymbol{\alpha} = (b_1, b_2, \dots, b_n)$$

$$\mathbf{T} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix}$$

$$\mathbf{t} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

para esta. Para probar este resultado tenemos que probar antes el siguiente lema.

Lema 4.2.1: Si la transformada de Laplace de la forma (4.2.7) corresponde a una distribución matriz exponencial entonces

$$\alpha_0 = 1 - \frac{b_1}{a_1}.$$

Demostración: Tenemos que

$$\mathbf{T}^{-1} = \begin{pmatrix} -\frac{a_2}{a_1} & -\frac{a_3}{a_1} & \dots & -\frac{a_n}{a_1} & -\frac{1}{a_1} \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

Utilizando que $\lim_{x \rightarrow 0^+} (1 + \boldsymbol{\alpha} \exp(\mathbf{T}x) \mathbf{T}^{-1} \mathbf{s}) = \alpha_0$ y la representación $(\boldsymbol{\alpha}, \mathbf{T}, \mathbf{s})$ tenemos que

$$\alpha_0 = \lim_{x \rightarrow 0^+} F(x) = 1 + \boldsymbol{\alpha} \mathbf{T}^{-1} \mathbf{s}$$

$$\alpha_0 = 1 - \frac{b_1}{a_1}$$

■

Lema 4.2.2: La inversa de $s\mathbf{I} - \mathbf{T}$, $\mathbf{A} = \{a_{ij}\}$, esté dada por

$$a_{1j} = \frac{s^{n-j} + a_1 s^{n-j-1} + \dots + a_{n-j-1} s + a_{n-j}}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}, \quad j = 1, 2, \dots, n-1$$

$$a_{1n} = \frac{1}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}$$

mientras que para $i > 1$,

$$a_{ij} = -\frac{a_{n+1-j} s^{i-2} + a_{n+2-j} s^{i-3} + \dots + a_n s^{i-j-1}}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}, \quad j = 1, 2, \dots, i-1$$

y para $j = i, i+1, \dots, n$

$$a_{ij} = \frac{s^{i-1}}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n} (s^{n-j} + a_1 s^{n-j-1} + \dots + a_{n-j})$$

Demostración: Consideremos solo el caso $\mathbf{A}(s\mathbf{I}-\mathbf{T})$ puesto que el caso $(s\mathbf{I}-\mathbf{T})\mathbf{A}$ es similar. Entonces el primer renglón de \mathbf{A} es diferente en estructura al resto de la matriz; empezaremos a revisar el producto entre el primer renglón de \mathbf{A} y la j -ésima columna de $(s\mathbf{I}-\mathbf{T})$

$$s\mathbf{I}-\mathbf{T} = \begin{pmatrix} s & 0 & \cdots & 0 \\ 0 & s & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix}$$

$$s\mathbf{I}-\mathbf{T} = \begin{pmatrix} s & -1 & 0 & \cdots & 0 & 0 \\ 0 & s & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s & -1 \\ -a_n & -a_{n-1} & -a_{n-2} & & a_2 & s+a_1 \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} s & -1 & 0 & \cdots & 0 & 0 \\ 0 & s & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s & -1 \\ -a_n & -a_{n-1} & -a_{n-2} & & a_2 & s+a_1 \end{pmatrix}$$

Si $j = 1$; existen solamente dos contribuciones a este producto

$$a_{11}s + a_{1n}a_n = \frac{s^{n-1} + a_1s^{n-2} + \cdots + a_{n-1}}{s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n} s + \frac{1}{s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n} a_n$$

$$a_{11}s + a_{1n}a_n = \frac{s^n + a_1s^{n-1} + \cdots + a_{n-1}s}{s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n} + \frac{a_n}{s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n}$$

$$a_{11}s + a_{1n}a_n = \frac{s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n}{s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n} = 1$$

$$a_{11}s + a_{1n}a_n = 1$$

Para $j > 1$; tenemos tres contribuciones

$$-a_{1,j-1} + a_{1j}s + a_{1n}a_{n-j+1} = -\frac{s^{n-j+1} + a_1s^{n-j} + \dots + a_{n-j+1}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} + \frac{s^{n-j} + a_1s^{n-j-1} + \dots + a_{n-j}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} s + \frac{1}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} a_{n-j+1}$$

$$-a_{1,j-1} + a_{1j}s + a_{1n}a_{n-j+1} = \frac{-s^{n-j+1} - a_1s^{n-j} - \dots - a_{n-j+1}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} + \frac{s^{n-j+1} + a_1s^{n-j} + \dots + sa_{n-j}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} + \frac{a_{n-j+1}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n}$$

$$-a_{1,j-1} + a_{1j}s + a_{1n}a_{n-j+1} = \frac{-s^{n-j+1} - a_1s^{n-j} - \dots - a_{n-j+1} + s^{n-j+1} + a_1s^{n-j} + \dots + a_{n-j+1}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n}$$

$$-a_{1,j-1} + a_{1j}s + a_{1n}a_{n-j+1} = 0$$

Ahora consideremos el producto del i -ésimo renglón de \mathbf{A} y la primera columna de $(s\mathbf{I} - \mathbf{T})$ para $i > 1$. Entonces las únicas contribuciones son de la primera y última entrada del renglón.

$$a_{i1}s + a_{in}a_n = -\frac{a_n s^{i-2}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} s + \frac{s^{i-1}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} a_n$$

$$a_{i1}s + a_{in}a_n = -\frac{a_n s^{i-1}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} + \frac{a_n s^{i-1}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n}$$

$$a_{i1}s + a_{in}a_n = 0$$

En seguida consideremos el producto del i -ésimo renglón de \mathbf{A} y la j -ésima columna $(s\mathbf{I} - \mathbf{T})$. Entonces existen tres contribuciones, una de -1 en la columna, una de s y una de a_{n+1-j} . Si $i < j$ tenemos que

$$-a_{i,j-1} + a_{ij}s + a_{in}a_{n-j+1} = -\frac{s^{i-1}(s^{n-j+1} + a_1s^{n-j} + \dots + a_{n-j+1})}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} + \frac{s^{i-1}(s^{n-j} + a_1s^{n-j-1} + \dots + a_{n-j})}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} s + \frac{s^{i-1}a_{n-j+1}}{s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n}$$

$$-a_{i,j-1} + a_{ij}s + a_{in}a_{n-j+1} = 0$$

Si $i = j$ la contribución es

$$-a_{i,i-1} + a_{ii}s + a_m a_{n-i+1} = \frac{s^{i-2} a_{n-i+2} + s^{i-3} a_{n-i+3} + \dots + a_n s^0}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n} +$$

$$\frac{s^{i-1} (s^{n-i} + a_1 s^{n-i-1} + \dots + a_{n-i})}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n} s + \frac{s^{i-1} a_{n-i+1}}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}$$

$$-a_{i,i-1} + a_{ii}s + a_m a_{n-i+1} = \frac{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n} = 1$$

Si $i \geq j+1$ tenemos que

$$-a_{i,j-1} + a_{ij}s + a_m a_{n-j+1} = \frac{s^{i-2} a_{n-j+2} + s^{i-3} a_{n-j+3} + \dots + a_n s^{i-j}}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n} -$$

$$\frac{s^{i-2} a_{n-j+1} + \dots + a_n s^{i-j-1}}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n} s + \frac{s^{i-1} a_{n-j+1}}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}$$

$$-a_{i,j-1} + a_{ij}s + a_m a_{n-j+1} = 0$$

■

Demostración de la Proposición 2.3: Primero demostraremos que la transformada de Laplace de una distribución matriz exponencial tiene la forma (4.2.7). Sea $(\boldsymbol{\beta}, \mathbf{S}, \mathbf{s})$ una representación de la distribución. Entonces la transformada de Laplace está dada por

$$\hat{f}(s) = \boldsymbol{\beta}(s\mathbf{I} - \mathbf{S})^{-1} \mathbf{s}.$$

Usando la descomposición de Jordan

$$\mathbf{S} = \mathbf{P} \begin{pmatrix} \mathbf{J}_1 & 0 & \dots & 0 & 0 \\ 0 & \mathbf{J}_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \mathbf{J}_{m-1} & 0 \\ 0 & 0 & \dots & 0 & \mathbf{J}_m \end{pmatrix} \mathbf{P}^{-1},$$

donde \mathbf{J}_i es una matriz de $n_i \times n_i$, $n_1 + n_2 + \dots + n_m = \dim(S)$, y

$$\mathbf{J}_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \vdots & \lambda_i & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix}$$

es la matriz de Jordan correspondiente al valor propio λ_i de \mathbf{S} con multiplicidad n_i . Entonces tenemos que

$$\mathbf{J}_i^{-1} = \begin{pmatrix} \frac{1}{\lambda_i} & -\frac{1}{\lambda_i^2} & \cdots & (-1)^{n_i} \frac{1}{\lambda_i^{n_i-1}} & (-1)^{n_i+1} \frac{1}{\lambda_i^{n_i}} \\ 0 & \frac{1}{\lambda_i} & \cdots & (-1)^{n_i-1} \frac{1}{\lambda_i^{n_i-2}} & (-1)^{n_i} \frac{1}{\lambda_i^{n_i-1}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_i} & -\frac{1}{\lambda_i^2} \\ 0 & 0 & \cdots & 0 & \frac{1}{\lambda_i} \end{pmatrix}$$

y

$$\begin{pmatrix} \mathbf{J}_1 & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{J}_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{J}_{m-1} & 0 \\ 0 & 0 & \cdots & 0 & \mathbf{J}_m \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{J}_1^{-1} & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{J}_2^{-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{J}_{m-1}^{-1} & 0 \\ 0 & 0 & \cdots & 0 & \mathbf{J}_m^{-1} \end{pmatrix}.$$

Sea $\tilde{\mathbf{J}}_i(s) = \mathbf{J}_i - s\mathbf{I}_{n_i}$. Entonces

$$\hat{f}(s) = \boldsymbol{\beta}(s\mathbf{I} - \mathbf{S})^{-1}\mathbf{s}$$

$$\hat{f}(s) = -\boldsymbol{\beta}\mathbf{P} \begin{pmatrix} \tilde{\mathbf{J}}_1^{-1} & 0 & \cdots & 0 & 0 \\ 0 & \tilde{\mathbf{J}}_2^{-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \tilde{\mathbf{J}}_{m-1}^{-1} & 0 \\ 0 & 0 & \cdots & 0 & \tilde{\mathbf{J}}_m^{-1} \end{pmatrix} \mathbf{P}^{-1}\mathbf{s}.$$

Pero

$$\tilde{\mathbf{J}}_i^{-1}(s) = \begin{pmatrix} \frac{1}{\lambda_i - s} & -\frac{1}{(\lambda_i - s)^2} & \cdots & (-1)^{n_i} \frac{1}{(\lambda_i - s)^{n_i-1}} & (-1)^{n_i+1} \frac{1}{(\lambda_i - s)^{n_i}} \\ 0 & \frac{1}{\lambda_i - s} & \cdots & (-1)^{n_i-1} \frac{1}{(\lambda_i - s)^{n_i-2}} & (-1)^{n_i} \frac{1}{(\lambda_i - s)^{n_i-1}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_i - s} & -\frac{1}{(\lambda_i - s)^2} \\ 0 & 0 & \cdots & 0 & \frac{1}{\lambda_i - s} \end{pmatrix},$$

introduciendo los vectores $\boldsymbol{\beta}' = -\boldsymbol{\beta}\mathbf{P}$ y $\mathbf{s}' = \mathbf{P}^{-1}\mathbf{s}$ en forma de bloque partido $\boldsymbol{\beta}' = (\boldsymbol{\beta}'_1, \dots, \boldsymbol{\beta}'_m)$ y $\mathbf{s}' = (\mathbf{s}'_1, \dots, \mathbf{s}'_m)$ podemos escribir

$$\hat{f}(s) = \sum_{i=1}^m \boldsymbol{\beta}'_i \tilde{\mathbf{J}}_i^{-1}(s) \mathbf{s}'_i,$$

lo cual muestra que $\hat{f}(s)$ puede en general ser rescrita como (4.2.7).

Ahora demostraremos que la transformada de Laplace (4.2.7) coincide con $\boldsymbol{\alpha}(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{t}$, donde $\boldsymbol{\alpha}$, \mathbf{T} y \mathbf{t} . Entonces \mathbf{t} tiene un 1 en la última posición y 0 en las demás, entonces $(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{t}$ es simplemente la columna derecha en la matriz $(s\mathbf{I} - \mathbf{T})^{-1}$. Pero por el Lema 4.2.1 tenemos que

$$a_m = \frac{s^{i-1}}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}$$

así

$$\boldsymbol{\alpha}(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{t} = \frac{\sum_{i=1}^n b_i s^{i-1}}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}.$$

■

El siguiente corolario muestra que, aún cuando la relación $\mathbf{t} = -\mathbf{T}\mathbf{e}$ no ocurre, siempre es posible elegir una representación en la cuál este caso ocurre.

Corolario 4.2.1: Para cualquier distribución matriz exponencial, existe una representación $(\boldsymbol{\alpha}, \mathbf{T}, \mathbf{t})$ tal que $\mathbf{t} = -\mathbf{T}\mathbf{e}$.

Debe haber una representación que tiene un orden más pequeño o mínimo. Una representación que tiene el orden más pequeño es llamada una representación mínima

Demostración: Consideremos la representación mínima $(\boldsymbol{\pi}, \mathbf{S}, \mathbf{s})$ de la distribución, y sea f la densidad de la distribución

$$f(x) = \boldsymbol{\pi} e^{\mathbf{S}x} \mathbf{s}$$

donde $\boldsymbol{\pi}$, \mathbf{S} , \mathbf{s} son como en la Proposición 4.2.1. Entonces tomamos una matriz no singular \mathbf{M} y escribimos

$$f(x) = \boldsymbol{\pi} \mathbf{M} e^{\mathbf{M}^{-1} \mathbf{S} \mathbf{M} x} \mathbf{M}^{-1} \mathbf{s}.$$

De aquí se sigue que el requisito es $\mathbf{M}^{-1} \mathbf{S} \mathbf{M} \mathbf{e} = -\mathbf{M}^{-1} \mathbf{s}$ o equivalentemente

$$\mathbf{M} \mathbf{e} = -\mathbf{S}^{-1} \mathbf{s}$$

Usando la estructura de \mathbf{S} por la proposición anterior observamos que las constantes en \mathbf{M} son tales que de los renglones 2 al n suman 0 mientras que el renglón 1 suma $\frac{1}{a_n}$. Por otra parte elegiremos \mathbf{M} de tal manera que sea no singular, una posible elección es

$$\mathbf{M} = \begin{pmatrix} \frac{1}{a_n} & 0 & 0 & 0 & \cdots & 0 & 0 \\ a_n & 0 & -1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

La matriz \mathbf{M} existe puesto que $a_n \neq 0$, el cual de hecho es el caso, porque la representación es mínima con $b_1 = a_n$ y $\hat{f}(\mathbf{0}) = 1$. Entonces sustituyendo $\mathbf{T} = \mathbf{M}^{-1} \mathbf{S} \mathbf{M}$, $\mathbf{t} = \mathbf{M}^{-1} \mathbf{s}$ y $\boldsymbol{\alpha} = \boldsymbol{\pi} \mathbf{M}$ se sigue el resultado. ■

Mientras que para las distribuciones tipo fase la cuestión de las representaciones mínimas es un problema abierto importante y difícil, veremos que en el caso matriz exponencial existe una solución simple. Por este motivo, necesitaremos la siguiente notación. Definimos

$$R_p = \text{span}\{\mathbf{s}, \mathbf{T}\mathbf{s}, \dots, \mathbf{T}^{p-1}\mathbf{s}\} \quad L_p = \text{span}\{\boldsymbol{\alpha}, \boldsymbol{\alpha}\mathbf{T}, \dots, \boldsymbol{\alpha}\mathbf{T}^{p-1}\}$$

$$R_\infty = \text{span}\{\mathbf{T}^i \mathbf{s} : i = 0, 1, 2, \dots\} \quad R_e = \text{span}\{e^{\mathbf{T}x} \mathbf{s} : x \geq 0\}$$

$$L_\infty = \text{span}\{\boldsymbol{\alpha}\mathbf{T}^i : i = 0, 1, 2, \dots\} \quad L_e = \text{span}\{\boldsymbol{\alpha} e^{\mathbf{T}x} : x \geq 0\}$$

Definición 4.2.1: La pareja (\mathbf{T}, \mathbf{s}) tiene la propiedad por la derecha $\dim(R_p) = p$. Análogamente, la pareja $(\boldsymbol{\alpha}, \mathbf{T})$ tiene la propiedad por la izquierda $\dim(L_p) = p$.

Equivalentemente, la propiedad por la derecha ocurre si $\mathbf{s}, \mathbf{T}\mathbf{s}, \dots, \mathbf{T}^{p-1}\mathbf{s}$ son linealmente independientes.

El principal resultado acerca de las representaciones mínimas es el siguiente:

Teorema 4.2.1: Una representación $(\boldsymbol{\alpha}, \mathbf{T}, \mathbf{s})$ de una distribución matriz-exponencial B es mínima si y sólo si la pareja (\mathbf{T}, \mathbf{s}) tiene la propiedad por la derecha y la pareja $(\boldsymbol{\alpha}, \mathbf{T})$ tiene la propiedad por la izquierda.

Lema 4.2.3: $R_p = R_\infty = R_e$.

Demostración: Obviamente $R_p \subseteq R_\infty$. Por el teorema de Cayley-Hamilton \mathbf{T}^p es una combinación lineal de $\mathbf{I}, \mathbf{T}, \dots, \mathbf{T}^{p-1}$. Por lo tanto también lo son $\mathbf{T}^{p+1}, \mathbf{T}^{p+2}, \dots$, de modo que $\mathbf{T}^i \mathbf{s} \in R_p$ para toda i (no solo para $i < p$), y por lo tanto $R_p = R_\infty$.

Consideremos la expansión en serie de $e^{\mathbf{T}x}\mathbf{s} = \sum_{n=0}^{\infty} \frac{x^n}{n!} \mathbf{T}^n \mathbf{s}$. Es obvio que $\sum_{n=0}^K \frac{x^n}{n!} \mathbf{T}^n \mathbf{s} \in R_\infty$ para toda K , entonces R_∞ es cerrado (es un vector de espacio finito dimensional), tenemos que $\lim_{K \rightarrow \infty} \sum_{n=0}^K \frac{x^n}{n!} \mathbf{T}^n \mathbf{s} \in R_\infty$. Así si $y \in R_e$ entonces $y = e^{\mathbf{T}x}\mathbf{s} = \sum_{n=0}^{\infty} \frac{x^n}{n!} \mathbf{T}^n \mathbf{s} \in R_\infty$ para alguna $x \geq 0$. Entonces $R_e \subseteq R_\infty$.

Para la inversa, notemos que también R_e es cerrado. Tomando $x = 0$ observamos que $\mathbf{s} \in R_e$. También $e^{\mathbf{T}x}\mathbf{s} \in R_e$ y por lo tanto también $\frac{1}{x}(e^{\mathbf{T}x}\mathbf{s} - \mathbf{s}) \in R_e$. Entonces R_e es cerrado y obtenemos que

$$\mathbf{T}\mathbf{s} = \lim_{x \downarrow 0} \frac{e^{\mathbf{T}x}\mathbf{s} - \mathbf{s}}{x} \in R_e$$

Igualmente

$$\mathbf{T}^2\mathbf{s} = \lim_{x \downarrow 0} 2 \frac{e^{\mathbf{T}x}\mathbf{s} - \mathbf{T}\mathbf{s}x - \mathbf{s}}{x^2} \in R_e$$

Para toda n tenemos que

$$\text{span}\{\mathbf{s}, \mathbf{T}\mathbf{s}, \dots, \mathbf{T}^n \mathbf{s}\} \subseteq R_e.$$

Por la propiedad de cerradura de R_e obtenemos que

$$R_\infty = \text{span}\{\mathbf{T}^n \mathbf{s} : n = 0, 1, 2, \dots\} \subseteq \bar{R}_e \subseteq R_e$$

(El caso $L_p = L_\infty = L_e$ es similar). ■

Proposición 4.2.4: Sean $(\boldsymbol{\alpha}, \mathbf{T}, \mathbf{s})$, $(\boldsymbol{\beta}, \mathbf{T}, \mathbf{s})$ dos representaciones de la misma distribución matriz exponencial. Si la pareja (\mathbf{T}, \mathbf{s}) tiene la propiedad derecha, entonces $\boldsymbol{\alpha} = \boldsymbol{\beta}$, mientras que si la propiedad derecha falla, entonces es posible obtener $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$.

Demostración: Supongamos que la propiedad derecha se cumple. Por igualdad de medidas, tenemos que $\boldsymbol{\alpha} e^{\mathbf{T}x} \mathbf{s} = \boldsymbol{\beta} e^{\mathbf{T}x} \mathbf{s}$ para toda $x \geq 0$. Por lo tanto $\boldsymbol{\alpha} - \boldsymbol{\beta}$ es ortogonal a $R_e = R_p = \mathbb{R}^p$, de modo que $\boldsymbol{\alpha} = \boldsymbol{\beta}$.

Supongamos ahora que la propiedad derecha falla. Entonces existe $\boldsymbol{\gamma} \neq 0$ tal que $\boldsymbol{\gamma}$ es ortogonal a R_p . Entonces $R_p = R_e$, $\boldsymbol{\gamma}$ es también ortogonal a cada $e^{\mathbf{T}x} \mathbf{s}$. Así, si $(\boldsymbol{\alpha}, \mathbf{T}, \mathbf{s})$ es una representación de la distribución matriz exponencial, también lo es $(\boldsymbol{\alpha} + \boldsymbol{\gamma}, \mathbf{T}, \mathbf{s})$. ■

Se puede probar que si $(\boldsymbol{\alpha}, \mathbf{T}, \mathbf{s})$ y $(\boldsymbol{\alpha}, \mathbf{T}, \mathbf{t})$ son dos representaciones de la misma distribución matriz exponencial entonces: si $(\boldsymbol{\alpha}, \mathbf{T})$ tiene la propiedad izquierda entonces $\mathbf{s} = \mathbf{t}$, mientras que si la propiedad izquierda falla es posible elegir $\mathbf{s} \neq \mathbf{t}$.

Proposición 4.2.5: Sea $(\boldsymbol{\alpha}, \mathbf{T}, \mathbf{s})$ una representación n -dimensional de una distribución matriz exponencial que no satisface la propiedad derecha. Sea $d_R = \dim(R_n) < n$ y sea $\mathbf{x}_1, \dots, \mathbf{x}_{d_R}$ una base para R_n . Definimos la matriz $\mathbf{A} = (\mathbf{x}_1, \dots, \mathbf{x}_{d_R})$ de $n \times d_R$ y definimos la transformación de cambio de base \mathbf{B} (matriz de $d_R \times n$) tal que $\mathbf{B}\mathbf{x}_i = \mathbf{e}_i$ para $i = 1, \dots, d_R$ donde $\mathbf{e}_1, \dots, \mathbf{e}_{d_R}$ denotan los vectores base Euclidianos. Entonces

$$(\boldsymbol{\alpha}\mathbf{A}, \mathbf{B}\mathbf{T}\mathbf{A}, \mathbf{B}\mathbf{s}) \tag{4.2.8}$$

es una representación para la distribución matriz exponencial de dimensión $d_R < n$. Similarmente, supongamos que la distribución no satisface la propiedad izquierda. Sea $d_L = \dim(L_n) < n$ y sea $\mathbf{y}_1, \dots, \mathbf{y}_{d_L}$ una base para L_n . Definimos la matriz de $d_L \times n$

$$\mathbf{C} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_{d_L} \end{pmatrix}$$

y \mathbf{D} (matriz de $n \times d_L$) tal que $\mathbf{y}_i \mathbf{D} = \mathbf{e}'_i$ para $i = 1, \dots, d_L$. Entonces

$$(\alpha \mathbf{D}, \mathbf{CTD}, \mathbf{Cs}) \quad (4.2.9)$$

es una representación de la distribución matriz exponencial de dimensión d_L

Demostración: Primero observamos que $\alpha \mathbf{B} \mathbf{y} = \mathbf{y}$ para toda $\mathbf{y} \in R_n$, y definimos

$$\tilde{\alpha} = \alpha \mathbf{A} \quad \tilde{\mathbf{T}} = \mathbf{B} \mathbf{T} \mathbf{A} \quad \tilde{\mathbf{s}} = \mathbf{B} \mathbf{s}$$

Entonces mostramos que $(\tilde{\alpha}, \tilde{\mathbf{T}}, \tilde{\mathbf{s}})$ es una representación de la distribución matriz exponencial

Esto se sigue por

$$\begin{aligned} \tilde{\alpha} \tilde{\mathbf{T}}^n \tilde{\mathbf{s}} &= \alpha (\mathbf{A} \mathbf{B} \mathbf{T})^n \mathbf{A} \mathbf{B} \mathbf{s} \\ \tilde{\alpha} \tilde{\mathbf{T}}^n \tilde{\mathbf{s}} &= \alpha (\mathbf{A} \mathbf{B} \mathbf{T})^n \mathbf{s} \quad (\text{entonces } \mathbf{s} \in R_n) \\ \tilde{\alpha} \tilde{\mathbf{T}}^n \tilde{\mathbf{s}} &= \alpha (\mathbf{A} \mathbf{B} \mathbf{T})^{n-1} (\mathbf{A} \mathbf{B}) (\mathbf{T} \mathbf{s}) \\ \tilde{\alpha} \tilde{\mathbf{T}}^n \tilde{\mathbf{s}} &= \alpha (\mathbf{A} \mathbf{B} \mathbf{T})^{n-1} \mathbf{T} \mathbf{s} \quad (\text{entonces } \mathbf{s} \in R_n) \\ &\vdots \\ \tilde{\alpha} \tilde{\mathbf{T}}^n \tilde{\mathbf{s}} &= \alpha \mathbf{T}^n \mathbf{s} \end{aligned}$$

y por lo tanto $\tilde{\alpha} e^{\tilde{\mathbf{T}}x} \tilde{\mathbf{s}} = \alpha e^{\mathbf{T}x} \mathbf{s}$ para toda x .

Esto ocurre de manera similar para

$$\tilde{\alpha} = \alpha \mathbf{D} \quad \tilde{\mathbf{T}} = \mathbf{C} \mathbf{T} \mathbf{D} \quad \tilde{\mathbf{s}} = \mathbf{B} \mathbf{s}$$

y de la misma forma $\tilde{\alpha} \tilde{\mathbf{T}}^n \tilde{\mathbf{s}} = \alpha \mathbf{T}^n \mathbf{s}$ y por lo tanto $\tilde{\alpha} e^{\tilde{\mathbf{T}}x} \tilde{\mathbf{s}} = \alpha e^{\mathbf{T}x} \mathbf{s}$ para toda x . ■

Definimos B_y como la distribución de $X - y$ dada $X > y$ donde X tiene distribución B .

Lema 4.2.3: Si B es matriz exponencial con representación $(\alpha, \mathbf{T}, \mathbf{s})$, entonces B_y es matriz exponencial con representación $(\alpha_y, \mathbf{T}, \mathbf{s})$ donde

$$\alpha_y = \frac{\alpha e^{\mathbf{T}y}}{\int_y^\infty \alpha e^{\mathbf{T}z} \mathbf{s} dz} = \frac{\alpha e^{\mathbf{T}y}}{\alpha e^{\mathbf{T}y} \mathbf{T}^{-1} \mathbf{s}}.$$

Demostración: En términos de la densidad, tenemos que

$$b_y(x) = \frac{b(x+y)}{1-B(y)} = \frac{\alpha e^{\mathbf{T}(x+y)} \mathbf{s}}{\int_y^\infty \alpha e^{\mathbf{T}z} \mathbf{s} dz} = \alpha_y e^{\mathbf{T}x} \mathbf{s}.$$

■

Demostración del Teorema 4.2.1: Se sigue de la Proposición 4.2.4 que una representación no puede ser minimal si la propiedad derecha falla. El caso donde la propiedad izquierda falla es simétrico, y así mostraremos que una representación que satisface las propiedades izquierda y derecha, es minimal.

Para una distribución matriz-exponencial B dada, sea p_0 el orden de la representación minimal, y sea q la dimensión de $V = \text{span}\{B_x : x \geq 0\}$, $q = \dim(V)$. Consideramos una representación $(\mathbf{a}, \mathbf{T}, \mathbf{s})$ de dimensión p . Se sigue por el Lema 4.2.2 que V es el espacio de medidas matriz exponencial con representación $(\boldsymbol{\beta}, \mathbf{T}, \mathbf{s})$ para alguna $\boldsymbol{\beta} \in L_e = L_p$. Por lo tanto $q \leq \dim(L_e) \leq p$. Además si la propiedad derecha ocurre, entonces $\boldsymbol{\beta}$ es único por la Proposición 4.2.2 tal que $q = \dim(L_e)$; si también la propiedad izquierda ocurre, entonces $\dim(L_e) = p$ y así $p = q$. Esto muestra que en cualquier caso, $q \leq p_0$, y que $q = p$ si las propiedades izquierda y derecha ocurren. Así, si las propiedades izquierda y derecha ocurren, $p \leq p_0$ tal que $p = p_0$ y la representación es minimal. ■

El siguiente corolario es consecuencia de la proposición anterior.

Corolario 4.2.3: La dimensión de la representación minimal es $d_R \wedge d_L$. Además, si $d_R < d_L$ es una representación minimal, y si $d_R > d_L$ es una representación minimal. Si $d_R = d_L$ entonces ambas son representaciones minimales.

Entonces una matriz \mathbf{T} que tiene forma diagonal significa lo siguiente

$$\mathbf{T} = \sum_{i=1}^p \lambda_i \mathbf{h}_i \mathbf{v}_i = \sum_{i=1}^p \lambda_i \mathbf{h}_i \otimes \mathbf{v}_i \quad (4.2.10)$$

donde las \mathbf{h}_i son vectores columna y las \mathbf{v}_i son vectores renglón (necesariamente vectores propios) que satisfacen $\mathbf{v}_i \mathbf{h}_j = \delta_{ij}$. Usando el Teorema 4.2.1 obtenemos el siguiente corolario.

Corolario 4.2.4: Supongamos que B es matriz exponencial con representación $(\mathbf{a}, \mathbf{T}, \mathbf{s})$ donde \mathbf{T} tiene la forma diagonal (4.2.10). Entonces la representación es minimal si y sólo si $\mathbf{a} \mathbf{h}_i \neq 0$ para toda i y $\mathbf{v}_j \mathbf{s} \neq 0$ para toda j .

Como veremos más adelante, la teoría de renovación nos conduce a expresiones donde \mathbf{T} es reemplazada por una matriz de la forma $(\mathbf{T} + \mathbf{s}\boldsymbol{\beta})$. Así el siguiente resultado es básico.

Proposición 4.2.6: Si la pareja (\mathbf{T}, \mathbf{s}) tiene la propiedad derecha, entonces también la tiene $(\mathbf{T} + \mathbf{s}\boldsymbol{\beta}, \mathbf{s})$ para cualquier vector renglón $\boldsymbol{\beta}$.

Demostración: Mostraremos por inducción que para $i = 0, \dots, p-1$

$$\text{span}\{\mathbf{s}, (\mathbf{T} + \mathbf{s}\boldsymbol{\beta})\mathbf{s}, \dots, (\mathbf{T} + \mathbf{s}\boldsymbol{\beta})^i \mathbf{s}\} = \text{span}\{\mathbf{s}, \mathbf{T}\mathbf{s}, \dots, \mathbf{T}^i \mathbf{s}\} \quad (4.2.11)$$

El caso para $i = 0$ es obvio. Suponiendo que es cierto para i , necesitamos mostrar que

$$(\mathbf{T} + \mathbf{s}\boldsymbol{\beta})^{i+1} \mathbf{s} \in \text{span}\{\mathbf{s}, \mathbf{T}\mathbf{s}, \dots, \mathbf{T}^{i+1}\mathbf{s}\}$$

Sin embargo, por la hipótesis de inducción podemos escribir $(\mathbf{T} + \mathbf{s}\boldsymbol{\beta})^i \mathbf{s} = \sum_{j=0}^i c_{ij} \mathbf{T}^j \mathbf{s}$ y así

$$(\mathbf{T} + \mathbf{s}\boldsymbol{\beta})^{i+1} \mathbf{s} = (\mathbf{T} + \mathbf{s}\boldsymbol{\beta})(\mathbf{T} + \mathbf{s}\boldsymbol{\beta})^i \mathbf{s}$$

$$(\mathbf{T} + \mathbf{s}\boldsymbol{\beta})^{i+1} \mathbf{s} = (\mathbf{T} + \mathbf{s}\boldsymbol{\beta}) \left(\sum_{j=0}^i c_{ij} \mathbf{T}^j \mathbf{s} \right)$$

$$(\mathbf{T} + \mathbf{s}\boldsymbol{\beta})^{i+1} \mathbf{s} = \sum_{j=0}^i c_{ij} \{ \mathbf{T}^{j+1} \mathbf{s} + \mathbf{s}(\boldsymbol{\beta} \mathbf{T}^j \mathbf{s}) \}$$

$$(\mathbf{T} + \mathbf{s}\boldsymbol{\beta})^{i+1} \mathbf{s} \in \text{span}\{\mathbf{s}, \mathbf{T}\mathbf{s}, \dots, \mathbf{T}^{i+1}\mathbf{s}\},$$

notemos que $\boldsymbol{\beta} \mathbf{T}^j \mathbf{s}$ es un número complejo. Así probamos que

$$\text{span}\{\mathbf{s}, (\mathbf{T} + \mathbf{s}\boldsymbol{\beta})\mathbf{s}, \dots, (\mathbf{T} + \mathbf{s}\boldsymbol{\beta})^i \mathbf{s}\} \subseteq \text{span}\{\mathbf{s}, \mathbf{T}\mathbf{s}, \dots, \mathbf{T}^i \mathbf{s}\}$$

La inclusión inversa se sigue reemplazando \mathbf{T} por $\mathbf{T} - \mathbf{s}\boldsymbol{\beta}$

■

Algunos ejemplos de distribuciones matriz exponenciales son:

Distribuciones Tipo Fase. Cada distribución tipo fase tiene una representación matriz exponencial $(\boldsymbol{\alpha}, \mathbf{T}, -\mathbf{T}\mathbf{e})$ donde $\boldsymbol{\alpha}$ es el vector de probabilidades de estados iniciales y \mathbf{T} es el generador infinitesimal de una cadena de Markov de estados finitos a tiempo continuo.

Distribuciones Hiperexponenciales Generalizadas. Tienen funciones de distribución, definidas para $u \geq 0$, de la forma

$$F(u) = \sum_{i=1}^n a_i (1 - e^{-\lambda_i u})$$

donde a_1, a_2, \dots, a_n son todos reales con $\sum_{i=1}^n a_i = 1$, y $\lambda_1 > \lambda_2 > \dots > \lambda_n > 0$.

Cada distribución hipereexponencial generalizada tiene una representación matriz exponencial $(\mathbf{a}, \mathbf{T}, \mathbf{t})$, donde

$$\mathbf{a} = (a_1 \quad a_2 \quad \cdots \quad a_n)$$

$$\mathbf{T} = \begin{pmatrix} -\lambda_1 & 0 & \cdots & 0 \\ 0 & -\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\lambda_n \end{pmatrix}$$

$$\mathbf{t} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Distribuciones cuyas transformadas de Laplace son el recíproco de un polinomio. Estas distribuciones tienen transformadas de Laplace racional por definición y son, por lo tanto, distribuciones matriz exponenciales.

4.3 Funciones de Distribución

Definición 4.3.1: Una función F es la función de distribución de una variable no negativa X si tiene las siguientes propiedades

$F(x)$ es no negativa para $x \geq 0$

$F(x)$ es no decreciente para $x \geq 0$; es decir, si $x < y$ entonces $F(x) < F(y)$

F es continua por la derecha; es decir, para $x \geq 0$, $\lim_{h \rightarrow 0^+} F(x+h) = F(x)$

$F(x) = 0$ para $x = 0$ y $\lim_{x \rightarrow \infty} F(x) = 1$

El siguiente lema simplifica la demostración del teorema de caracterización para distribuciones matriz exponenciales.

Lema 4.3.1: Sean \mathbf{a} y \mathbf{b} definidos de la siguiente manera:

$$\mathbf{a} = (a_1 \quad a_2 \quad \cdots \quad a_n)$$

$$\mathbf{b} = (b_1 \quad b_2 \quad \cdots \quad b_n)$$

Sea

$$a(s) = s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n$$

y

$$\mathbf{T} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix}. \quad (4.3.1)$$

Si, para $x > 0$

$$f(x) = \mathbf{b} \exp(\mathbf{T}x)\mathbf{s} \geq 0 \quad (4.3.2)$$

y

$$\int_0^\infty f(x)dx < \infty, \quad (4.3.3)$$

entonces existe una raíz de $a(s)$ de la parte real que es real y negativa.

Demostración: Sean $\xi_1, \xi_2, \dots, \xi_q$ las distintas raíces de $a(s)$ con sus respectivas multiplicidades n_1, n_2, \dots, n_q . Supongamos sin pérdida de generalidad que $\Re(\xi_1) \leq \Re(\xi_2) \leq \dots \leq \Re(\xi_q)$. Supongamos por contradicción que no existe una de $a(s)$ de la parte real que sea real. Entonces existirá r con $1 \leq r \leq \left\lfloor \frac{q}{2} \right\rfloor$, pares complejos conjugados que son las raíces de la parte real. Para $j = q - 2r + 1, q - 2r + 2, \dots, q$, están dados por

$$\xi_j = \gamma + i\beta_j$$

donde γ y $\beta_j \neq 0$ son reales. Notemos que para $j = q - 2r + 1, q - 2r + 2, \dots, q - 1$, $\beta_j = -\beta_{j+1} > 0$ sin pérdida de generalidad, y todos los β_j 's son distintos. Las raíces restantes $\xi_1, \xi_2, \dots, \xi_{q-2r}$, tienen parte real menor que γ .

Entonces las raíces de $a(s)$ son idénticos a los valores propios de \mathbf{T} ; por lo tanto, tenemos que

$$f(u) = \sum_{j=q-2r+1}^q P_j(u) e^{(\gamma+i\beta_j)u} + \sum_{j=1}^{q-2r} P_j(u) e^{\xi_j u}$$

donde para $j = 1, 2, \dots, q$, $P_j(u)$ son todos los posibles polinomios complejos de grado $n_j - 1$.

Sea $n = \max(n_{q-2r+1}, n_{q-2r+2}, \dots, n_q) - 1$. Consideremos

$$u^{-n} e^{-\gamma u} f(u) = \sum_{j=q-2r+1}^q u^{-n} P_j(u) e^{i\beta_j u} + \sum_{j=1}^{q-2r} u^{-n} P_j(u) e^{-(\gamma - \xi_j)u}$$

Ahora para cualquier $\varepsilon_1 > 0$, existe una $K_1 > 0$ tal que, para $u > K_1$,

$$\left| \sum_{j=1}^{q-2r} u^{-n} P_j(u) e^{-(\gamma - \xi_j)u} \right| < \varepsilon_1 \quad (4.3.4)$$

entonces, para $j = 1, 2, \dots, q-2r$, $\gamma > \Re(\xi_j)$. También, para cualquier $\varepsilon_2 > 0$, existe una $K_2 > 0$ tal que, para $u > K_2$,

$$\left| \sum_{j=q-2r+1}^q u^{-n} P_j(u) e^{i\beta_j u} - \sum_{j=q-2r+1}^q A_j e^{i\beta_j u} \right| < \varepsilon_2 \quad (4.3.5)$$

donde, para $j = q-2r+1, q-2r+2, \dots, q$, A_j es el coeficiente de u^n en $P_j(u)$. Ahora consideremos

$$\sum_{j=q-2r+1}^q A_j e^{i\beta_j u} = \sum_{\substack{j=q-2r+1 \\ \text{saltos de 2}}}^{q-1} B_j \cos \beta_j u + C_j \sen \beta_j u \quad (4.3.6)$$

donde, para $j = q-2r+1, q-2r+3, \dots, q-1$, B_j y C_j son reales. Consideremos la situación

donde para $j = q-2r+1, q-2r+3, \dots, q-1$, β_j es racional, esto es $\beta_j = \frac{r_j}{s_j}$ donde r_j y s_j

son enteros. Ahora para cualquier $u_1 > \max(K_1, K_2)$, existe un $u_2 > u_1$ tal que

$$\int_{u_1}^{u_2} \left(\sum_{\substack{j=q-2r+1 \\ \text{saltos de 2}}}^q B_j \cos \beta_j u + C_j \sen \beta_j u \right) du = 0 \quad (4.3.7)$$

como el integrando es periódico. Eligiendo

$$u_2 = u_1 + s_{q-2r+1} s_{q-2r+3} \dots s_{q-1} \times 2\pi$$

es suficiente. Entonces (4.3.6) no es idénticamente cero para $u_1 < u < u_2$; las funciones trigonométricas son todas linealmente independientes porque los β_j 's son todos distintos y al menos uno de los B_j 's y C_j 's es distinto a cero entonces existe alguna $u > \max(K_1, K_2)$ con $u_1 < u < u_2$ y $\varepsilon_3 > 0$ tal que

$$\sum_{\substack{j=q-2r+1 \\ \text{saltos de 2}}}^q B_j \cos \beta_j u + C_j \sen \beta_j u < -\varepsilon_3 \quad (4.3.8)$$

Si cualquier β_j es un número irracional aproximando es arbitrariamente cercano por un número racional de la forma (4.3.8).

Ahora elegimos ε_1 y ε_2 tal que $\varepsilon_1 + \varepsilon_2 < \varepsilon_3$. Esto puede ser logrado porque (4.3.7) ocurre para cualquier $u_1 > \max(K_1, K_2)$. Ahora tenemos una $u_1 > \max(K_1, K_2)$, con $u_1 < u < u_2$, tal que

$$u^{-n} e^{-\gamma u} f(u) = \sum_{j=q-2r+1}^q u^{-n} P_j(u) e^{i\beta_j u} + \sum_{j=1}^{q-2r} u^{-n} P_j(u) e^{-(\gamma-\xi_j)u}$$

$$u^{-n} e^{-\gamma u} f(u) < \sum_{j=q-2r+1}^q A_j e^{i\beta_j u} + \varepsilon_2 + \varepsilon_1$$

$$u^{-n} e^{-\gamma u} f(u) < \varepsilon_3 + \varepsilon_2 + \varepsilon_1$$

$$u^{-n} e^{-\gamma u} f(u) < 0$$

La primera desigualdad ocurre de (4.3.4) y (4.3.5), y la segunda de (4.3.8). Así existe una $u > 0$ tal que $f(u) < 0$. Esto contradice (4.3.2), por lo tanto existe un cero de $a(s)$ de la máxima parte real, que es real.

Supongamos que existe un cero $a(s)$ de la máxima parte real, que es real y no negativa. Sin pérdida de generalidad, sean estos ceros $\xi_q = \sigma \geq 0$. El polinomio $a(s)$ debe tener un número de ceros en los pares complejos conjugados cuya parte real es igual a σ . Sean el número de pares conjugados que son distintos de s , donde $0 \leq s \leq \left\lfloor \frac{q-1}{2} \right\rfloor$. Esto es, para $j = q-2s, q-2s+1, \dots, q-1$,

$$\xi_j = \sigma + i\eta_j,$$

donde $\eta_j \neq 0$ es real. Notemos que para $j = q-2s, q-2s+2, \dots, q-2$, $\eta_j = -\eta_{j+1} > 0$ sin pérdida de generalidad. Los ceros restantes $\xi_1, \xi_2, \dots, \xi_{q-2s-1}$, todos tiene partes reales menores que σ . Tenemos que

$$g(u) = Q_q(u) e^{\sigma u} + \sum_{j=q-2s}^{q-1} Q_j(u) e^{(\sigma+\eta_j)u} + \sum_{j=1}^{q-2s-1} Q_j(u) e^{\xi_j u}, \quad (4.3.9)$$

donde para $j = 1, 2, \dots, q$, $Q_j(u)$ es posiblemente un polinomio complejo de grado $\eta_j - 1$. Si $\sigma > 0$ entonces (4.3.9) diverge conforme $u \rightarrow \infty$. Si $\sigma = 0$ entonces (4.3.9) puede divergir, oscilar o aproximarse a una constante distinta de cero conforme $u \rightarrow \infty$. En cualquier caso (4.3.3) no se satisface. Esto es una contradicción y por lo tanto cualquier cero de $a(s)$ de la máxima parte real debe ser también negativa. ■

Teorema 4.3.1: Sea $f(u)$, definida por (4.3.2) tal que (4.3.3) ocurre. Los vectores \mathbf{a} y \mathbf{b} corresponden a una distribución matriz exponencial no trivial si y solo si

$$f(u) \geq 0 \text{ para } u > 0$$

$$0 < \frac{b_1}{a_1} < 1$$

existe un cero de $a(s)$ de la máxima parte real tal que es real y negativa.

Demostración: Supongamos que los vectores \mathbf{a} y \mathbf{b} corresponden a una distribución matriz exponencial no trivial. Si \mathbf{T} es definida por (4.3.1) entonces la distribución tiene una representación (\mathbf{a}, \mathbf{T}) . La correspondiente función de distribución $F(u)$, la cuál satisface las condiciones de la Definición 4.3.1 es de la forma (4.2.1). Tenemos que

$$F(u) = \int_0^u f(t)dt + \alpha_0 \quad (4.3.10)$$

donde $0 \leq \alpha_0 \leq 1$. Entonces $F(u)$ es no decreciente para $u \geq 0$, $f(u) \geq 0$ para $u > 0$.

4.4 Teoría de Renovación

Consideraremos un proceso de renovación donde la distribución entre arribos \mathbf{B} es matriz exponencial con representación $(\mathbf{a}, \mathbf{T}, \mathbf{s})$.

Teorema 4.4.1: La densidad de renovación está dada por

$$u(t) = \mathbf{a}e^{(\mathbf{T} + \mathbf{s}\mathbf{a})t} \mathbf{s} \quad (4.4.1)$$

Demostración del Teorema 4.4.1 para distribuciones tipo fase: Consideremos el proceso de Markov por pedazos de los procesos de Markov terminales en los cuales los tiempos de vida son los tiempos entre los arribos. Entonces existen dos tipos de saltos en este proceso, los saltos en los procesos de Markov terminales y los saltos que surgen de los arribos. Entonces la transición del estado i al j toma lugar de dos maneras, cualquiera dentro del proceso de Markov terminal el cuál nos lleva a una intensidad t_{ij} , o si un proceso de Markov termina en el estado i y el siguiente proceso de Markov inicia en el estado j el cuál nos lleva a una intensidad de $t_i\alpha_j$. Por lo tanto la intensidad de transición del estado i al estado j es simplemente la suma de las intensidades $t_{ij} + t_i\alpha_j$. Por lo tanto la matriz de intensidad está dada por $\mathbf{T} + \mathbf{s}\mathbf{a}$, de ahí la matriz de transición está dada por $P^t = \exp\{(\mathbf{T} + \mathbf{s}\mathbf{a})t\}$. Ahora, la densidad de renovación tiene la interpretación $u(t)dt =$ probabilidad de una renovación en $[t, t + dt)$, entonces de acuerdo con la ley de la probabilidad total

$$u(t)dt = \sum_{j=1}^p \alpha_j p_{ji}^t t_i dt$$

entonces $t_i dt$ es la probabilidad de salir en el estado i en el tiempo dt . ■

En el caso de las distribuciones matriz-exponencial ninguna interpretación probabilística parece inherente y esto puede ser por lo tanto inesperado. Sin embargo la prueba es más formal y está basada en el siguiente lema.

$$\text{Lema 4.4.1: } \beta(-s\mathbf{I} - \mathbf{T} - s\boldsymbol{\alpha})^{-1}\mathbf{s} = \frac{\beta(-s\mathbf{I} - \mathbf{T})^{-1}\mathbf{s}}{1 - \boldsymbol{\alpha}(-s\mathbf{I} - \mathbf{T})^{-1}\mathbf{s}}$$

Demostración:

$$(\mathbf{A} + \mathbf{UBV})^{-1} = \mathbf{A}^{-1}\mathbf{UB}(\mathbf{B} + \mathbf{BVA}^{-1}\mathbf{UB})^{-1}\mathbf{BVA}^{-1}$$

con $\mathbf{A} = -s\mathbf{I} - \mathbf{T}$, $\mathbf{U} = -\mathbf{s}$, $\mathbf{B} = \mathbf{I}$ y $\mathbf{V} = \boldsymbol{\alpha}$. Tenemos que

$$\beta(-s\mathbf{I} - \mathbf{T} - s\boldsymbol{\alpha})^{-1}\mathbf{s} = \beta(\mathbf{A} - s\boldsymbol{\alpha})^{-1}\mathbf{s}$$

$$\beta(-s\mathbf{I} - \mathbf{T} - s\boldsymbol{\alpha})^{-1}\mathbf{s} = \beta\left\{\mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{s}(\mathbf{I} - \boldsymbol{\alpha}\mathbf{A}^{-1}\mathbf{s})^{-1}\boldsymbol{\alpha}\mathbf{A}^{-1}\right\}\mathbf{s}$$

$$\beta(-s\mathbf{I} - \mathbf{T} - s\boldsymbol{\alpha})^{-1}\mathbf{s} = \gamma_2 + \frac{\gamma_1\gamma_2}{1 - \gamma_1}$$

$$\beta(-s\mathbf{I} - \mathbf{T} - s\boldsymbol{\alpha})^{-1}\mathbf{s} = \frac{\gamma_2 - \gamma_1\gamma_2 + \gamma_1\gamma_2}{1 - \gamma_1}$$

$$\beta(-s\mathbf{I} - \mathbf{T} - s\boldsymbol{\alpha})^{-1}\mathbf{s} = \frac{\gamma_2}{1 - \gamma_1}$$

donde $\gamma_2 = \beta\mathbf{A}^{-1}\mathbf{s}$ y $\gamma_1 = \boldsymbol{\alpha}\mathbf{A}^{-1}\mathbf{s}$

Demostración del Teorema 4.4.1. Por la definición de u es claro que la transformada es

$$\hat{u}(s) = \sum_{n=1}^{\infty} \hat{b}^n(s) = \frac{\hat{b}(s)}{1 - \hat{b}(s)} = \frac{\boldsymbol{\alpha}(-s\mathbf{I} - \mathbf{T})^{-1}\mathbf{s}}{1 - \boldsymbol{\alpha}(-s\mathbf{I} - \mathbf{T})^{-1}\mathbf{s}}$$

También la transformada de (4.4.1) es

$$\hat{u}^n(s) = \boldsymbol{\alpha}(-s\mathbf{I} - \mathbf{T} - s\boldsymbol{\alpha})^{-1}\mathbf{s}$$

(las funciones están definidas al menos para $\Re s < 0$). Por el Lema 4.4.1 con $\beta = \boldsymbol{\alpha}$ se sigue que $\hat{u} = \hat{u}^*$. Por lo tanto $u = u^*$. ■

Ahora encontraremos la densidad de renovación de un proceso de renovación retrasado.

Corolario 4.4.1: Consideremos un proceso de renovación retrasado con $b_0(x) = \beta e^{\mathbf{T}x} \mathbf{s}$ y $b(x) = \alpha e^{\mathbf{T}x} \mathbf{s}$, donde b_0 es la densidad retrasada y b es la densidad entre los arribos. Entonces la densidad de renovación del proceso de renovación retrasado está dado por:

$$u_0(x) = \beta e^{(\mathbf{T} + \alpha \mathbf{s})x} \mathbf{s}.$$

Demostración: La transformada de un proceso retrasado está dada por

$$\hat{u}_0 = \sum_{i=1}^{\infty} \hat{b}_0 \hat{b}^{i-1} = \frac{\hat{b}_0}{1 - \hat{b}}$$

y usando el lema anterior junto con la transformada de Laplace de $\beta e^{(\mathbf{T} + \alpha \mathbf{s})x} \mathbf{s}$ obtenemos el resultado. ■

Sea $\xi(t)$ el exceso de vida al tiempo t , i.e. al tiempo de espera hasta el siguiente arribo.

Teorema 4.4.2: $\xi(t)$ es matriz exponencial con representación $(\beta_t, \mathbf{T}, \mathbf{s})$ donde $\beta_t = \alpha e^{(\mathbf{T} + \alpha \mathbf{s})t}$. Una representación alternativa es $(\alpha, \mathbf{T}, e^{(\mathbf{T} + \alpha \mathbf{s})t} \mathbf{s})$. En el caso retrasado $b_0(x) = \beta e^{\mathbf{T}x} \mathbf{s}$, el mismo resultado ocurre con $\beta_t = \beta e^{(\mathbf{T} + \alpha \mathbf{s})t}$.

Demostración: Primero demostraremos el Teorema 4.4.2. para las distribuciones tipo fase. Consideremos el proceso de Markov en el eje real positivo el cual está dado juntando los procesos de Markov de saltos terminales. Entonces, la matriz de transición del proceso es $\exp((\mathbf{T} + \alpha \mathbf{s})t)$, así la distribución del proceso al tiempo t es $\beta_t = \beta e^{(\mathbf{T} + \alpha \mathbf{s})t}$ ($\beta = \alpha$ en el caso no retrasado), y por lo tanto esta es la distribución inicial para un proceso de Markov terminal con matriz de intensidad \mathbf{T} .

Ahora demostraremos el Teorema 4.4.2 para el caso matriz exponencial. En el caso no retrasado, la distribución del exceso de tiempo de vida es fácilmente vista que cumpla la ecuación de renovación.

$$\mathbb{P}(\xi(t) \leq z) = B(t+z) - B(t) + \int_0^t \mathbb{P}(\xi(t-x) \leq z) b(x) dx$$

la cuál tiene una solución única

$$\mathbb{P}(\xi(t) \leq z) = B(t+z) - B(t) + \int_0^t (B(t+z-x) - B(t-x)) u(x) dx.$$

Tomando la derivada con respecto a z nos lleva a la densidad $f_{\xi(t)}(z)$ de $\xi(t)$ como

$$f_{\xi(t)}(z) = b(t+z) + \int_0^t u(x) b(t+z-x) dx \quad (4.4.2)$$

ahora la integral es

$$\int_0^t \mathbf{a} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} \mathbf{s} \mathbf{a} e^{\mathbf{T}(t+z-x)} \mathbf{s} dx$$

notemos que

$$\frac{d}{dx} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} = e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} (\mathbf{T} + \mathbf{s}\mathbf{a})$$

entonces

$$e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} \mathbf{s}\mathbf{a} = \frac{d}{dx} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} - e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} \mathbf{T}$$

Insertando esto y usando integración parcial

$$f_{\xi(t)}(z) = b(t+z) + \int_0^t u(x) b(t+z-x) dx$$

$$f_{\xi(t)}(z) = b(t+z) + \int_0^t \mathbf{a} \left\{ \frac{d}{dx} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} - e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} \mathbf{T} \right\} e^{\mathbf{T}(t+z-x)} \mathbf{s} dx$$

$$f_{\xi(t)}(z) = b(t+z) + \int_0^t \mathbf{a} \frac{d}{dx} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} e^{\mathbf{T}(t+z-x)} \mathbf{s} dx - \int_0^t \mathbf{a} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} \mathbf{T} e^{\mathbf{T}(t+z-x)} \mathbf{s} dx$$

$$f_{\xi(t)}(z) = b(t+z) + \left[\mathbf{a} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} e^{\mathbf{T}(t+z-x)} \mathbf{s} \right]_0^t + \int_0^t \mathbf{a} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} \mathbf{T} e^{\mathbf{T}(t+z-x)} dx - \int_0^t \mathbf{a} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} \mathbf{T} e^{\mathbf{T}(t+z-x)} \mathbf{s} dx$$

$$f_{\xi(t)}(z) = b(t+z) + \mathbf{a} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})t} e^{\mathbf{T}_z} \mathbf{s} - b(t+z)$$

$$f_{\xi(t)}(z) = \mathbf{a} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})t} e^{\mathbf{T}_z} \mathbf{s}$$

Mostrando la representación alternativa, en lugar de escribir

$$f_{\xi(t)}(z) = b(t+z) + \int_0^t u(x) b(t+z-x) dx$$

escribimos

$$f_{\xi(t)}(z) = b(t+z) + \int_0^t b(t+z-x) u(x) dx$$

y usamos esto también

$$\frac{d}{dx} e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x} = (\mathbf{T} + \mathbf{s}\mathbf{a}) e^{(\mathbf{T}+\mathbf{s}\mathbf{a})x}$$

Finalmente, en el caso retrasado condicionando antes de la última renovación antes del tiempo t obtenemos

$$f_{\xi(t)}(z) = b_0(t+z) + \int_0^t u(x)b(t+z-x)dx.$$

El resto de los cálculos son similares que en el caso no retrasado. ■

Una distribución matriz-exponencial B con representación $(\alpha, \mathbf{T}, \mathbf{s})$ es defectuosa si la masa total $\|B\| = -\alpha\mathbf{T}^{-1}\mathbf{s}$ es menor que uno. Usualmente, esta es interpretada como una distribución que tiene un átomo con masa $1 - \|B\| = 1 + \alpha\mathbf{T}^{-1}\mathbf{s}$ en ∞ . Si esta masa perdida es reemplazada por 0, llamaremos a la distribución cero-modificada.

Corolario 4.4.2: El tiempo de vida ζ de un proceso de renovación con distribución matriz-exponencial defectuosa B con representación $(\alpha, \mathbf{T}, \mathbf{s})$ es matriz-exponencial cero-modificada con representación $(\alpha, \mathbf{T} + \mathbf{s}\alpha, (1 + \alpha\mathbf{T}^{-1}\mathbf{s})\mathbf{s})$.

Demostración: En términos de la función de distribución G y de la función de densidad g , tenemos

$$G(x) = \mathbb{P}(\zeta > x) = \mathbb{P}(\zeta(x) < \infty) = \int_0^\infty \alpha e^{(\mathbf{T} + \mathbf{s}\alpha)x} e^{\mathbf{T}y} \mathbf{s} dy = -\alpha e^{(\mathbf{T} + \mathbf{s}\alpha)x} \mathbf{T}^{-1} \mathbf{s}$$

$$g(x) = \alpha e^{(\mathbf{T} + \mathbf{s}\alpha)x} (\mathbf{T} + \mathbf{s}\alpha) \mathbf{T}^{-1} \mathbf{s} = \alpha e^{(\mathbf{T} + \mathbf{s}\alpha)x} (1 + \alpha\mathbf{T}^{-1}\mathbf{s}) \mathbf{s}.$$

■

4.4.1 Diagonalización

Asmussen y Bladt (1997) presentaron un método el cual permite escribir $\mathbf{T} + \mathbf{s}\alpha$ de forma diagonal y directamente obtener la función matriz exponencial necesitada para la densidad de renovación.

La herramienta básica son las raíces de la ecuación $\hat{B}[-s] = 1$. Escribimos $\hat{B}[s] = \frac{q(s)}{r(s)}$ donde

$q(s)$, $r(s)$ son polinomios sin raíces comunes y extendiendo el dominio de definición al plano complejo entero con las raíces de $r(s)$ eliminadas.

Por conveniencia supondremos que la representación es minimal.

Lema 4.4.1.1: Si $(\alpha, \mathbf{T}, \mathbf{s})$ es la representación minimal de B , entonces ningún valor propio λ de \mathbf{T} puede ser solución de $\hat{B}[-\lambda] = 1$.

Demostración: Supongamos que $(\mathbf{a}, \mathbf{T}, \mathbf{t})$ es una representación minimal. Además supongamos que λ satisface $\hat{B}[-\lambda] = 1$. Tenemos que demostrar que λ no es un valor propio de \mathbf{T} . Supongamos lo contrario. Haciendo la descomposición de Jordan como en la demostración de la Proposición 4.2.3 podemos escribir

$$\hat{B}[-\gamma] = \sum_{i=1}^m \alpha'_i \tilde{\mathbf{J}}_i^{-1}(\gamma) \mathbf{t}'_i$$

donde $\tilde{\mathbf{J}}_i(s) = \mathbf{J}_i - s\mathbf{I}_{n_i}$.

Por continuidad en una vecindad de λ tenemos que $\hat{B}[-\gamma] \rightarrow 1$ conforme $\gamma \rightarrow \lambda$. Pero entonces λ es un valor propio para \mathbf{T} , digamos λ_j , entonces conforme $\gamma \rightarrow \lambda$ todas las entradas triangulares superiores de $\tilde{\mathbf{J}}_j^{-1}(\gamma)$ convergen a ∞ . Así debemos tener que $\alpha'_i \tilde{\mathbf{J}}_j^{-1}(\gamma) \mathbf{t}'_i = 0$, pero entonces podemos escribir también

$$\hat{B}[-\gamma] = \sum_{i \neq j}^m \alpha'_i \tilde{\mathbf{J}}_i^{-1}(\gamma) \mathbf{t}'_i$$

de la cuál se sigue que la dimensión de esta representación es menor que la representación original $(\mathbf{a}, \mathbf{T}, \mathbf{s})$, pero esto es imposible porque por hipótesis $(\mathbf{a}, \mathbf{T}, \mathbf{s})$ es la representación minimal. ■

Proposición 4.4.1.1: Supongamos que $(\mathbf{a}, \mathbf{T}, \mathbf{s})$ es una representación minimal. Entonces λ es el valor propio para la matriz $\mathbf{T} + \mathbf{s}\mathbf{a}$ si y sólo si $\hat{B}[-\lambda] = 1$. En este caso, los correspondientes vectores propios izquierdo y derecho están dados respectivamente por:

$$\mathbf{r} = (\mathbf{T} - \lambda\mathbf{I})^{-1} \mathbf{s} \quad (4.4.1.1)$$

$$\mathbf{l} = \mathbf{a}(\mathbf{T} - \lambda\mathbf{I})^{-1}. \quad (4.4.1.2)$$

Demostración: Considere la ecuación $(\mathbf{T} + \mathbf{s}\mathbf{a})\mathbf{h} = \lambda\mathbf{h}$. Entonces $(\mathbf{T} - \lambda\mathbf{I})\mathbf{h} = -\mathbf{s}\mathbf{a}\mathbf{h} = -(\mathbf{a} \cdot \mathbf{h})\mathbf{s}$ donde “ \cdot ” denota el producto escalar. Sea $\mathbf{c}_k = -\mathbf{a} \cdot \mathbf{h}$. Entonces

$$\lambda\mathbf{h} = (\mathbf{T} + \mathbf{s}\mathbf{a})\mathbf{h}$$

$$\lambda\mathbf{h} = (\mathbf{T} + \mathbf{s}\mathbf{a})(\mathbf{T} - \lambda\mathbf{I})^{-1} \mathbf{c}_k \mathbf{s}$$

$$\lambda\mathbf{h} = (\mathbf{T} - \lambda\mathbf{I} + \lambda\mathbf{I} + \mathbf{s}\mathbf{a})(\mathbf{T} - \lambda\mathbf{I})^{-1} \mathbf{c}_k \mathbf{s}$$

$$\lambda\mathbf{h} = \left\{ \mathbf{I} + \lambda(\mathbf{T} - \lambda\mathbf{I})^{-1} + \mathbf{s}\mathbf{a}(\mathbf{T} - \lambda\mathbf{I})^{-1} \right\} \mathbf{c}_k \mathbf{s}$$

$$\lambda\mathbf{h} = \mathbf{c}_k \mathbf{s} + \lambda\mathbf{h} + \mathbf{s}\mathbf{a}(\mathbf{T} - \lambda\mathbf{I})^{-1} \mathbf{c}_k \mathbf{s}$$

que nos da

$$\mathbf{s}\alpha(\mathbf{T} - \lambda\mathbf{I})^{-1}\mathbf{c}_k\mathbf{s} = -\mathbf{c}_k\mathbf{s}.$$

Entonces $c_k \neq 0$ (por minimalidad), esta ecuación es equivalente a

$$\alpha(\mathbf{T} - \lambda\mathbf{I})^{-1}\mathbf{s} = -1.$$

Cálculos similares nos muestran que los vectores propios sugeridos son en verdad los vectores propios. ■

El siguiente teorema da una fórmula general para la densidad de renovación.

Proposición 4.4.1.2: Supongamos que la representación es minimal y la ecuación $\hat{B}[-s] = 1$ tiene p distintas soluciones $\lambda_1, \lambda_2, \dots, \lambda_p$ y sean $\mathbf{r}_k, \mathbf{l}_k$ definidos como en (4.4.1.1) y (4.4.1.2) con $\lambda = \lambda_k$. Entonces la densidad de renovación puede ser escrita como

$$u(x) = \sum_{k=1}^p c_k e^{\lambda_k x}, \text{ donde } c_k = \frac{(\alpha \cdot \mathbf{r}_k)(\mathbf{l}_k \cdot \mathbf{s})}{\mathbf{l}_k \cdot \mathbf{r}_k}$$

Demostración: Se sigue por las identidades de la diagonalización estándar

$$\mathbf{T} + \mathbf{s}\alpha = \sum_{k=1}^p \lambda_k \frac{\mathbf{r}_k \mathbf{l}_k}{\mathbf{l}_k \cdot \mathbf{r}_k}, \quad e^{\mathbf{T} + \mathbf{s}\alpha} = \sum_{k=1}^p e^{\lambda_k} \frac{\mathbf{r}_k \mathbf{l}_k}{\mathbf{l}_k \cdot \mathbf{r}_k}$$

Esta proposición tiene algunas consecuencias inmediatas para la función de renovación $U(x) = \mathbb{E}(N_t) = 1 + \int_0^x u(t) dt$ y la varianza $Var(N_t)$ donde N_t es el número de renovaciones hasta el tiempo t .

Proposición 4.4.1.3: Supongamos que la representación es minimal y la ecuación $\hat{B}[-s] = 1$ tiene p distintas soluciones $\lambda_1, \lambda_2, \dots, \lambda_p$, y sean $\mathbf{r}_k, \mathbf{l}_k$ definidos de la siguiente manera:

$$\mathbf{r} = (\mathbf{T} - \lambda\mathbf{I})^{-1}\mathbf{s}$$

$$\mathbf{l} = \alpha(\mathbf{T} - \lambda\mathbf{I})^{-1}$$

con $\lambda = \lambda_k$. Entonces la densidad de renovación puede ser escrita como

$$u(x) = \sum_{k=1}^p c_k e^{\lambda_k x}, \text{ donde } c_k = \frac{(\alpha \cdot \mathbf{r}_k)(\mathbf{l}_k \cdot \mathbf{s})}{\mathbf{l}_k \cdot \mathbf{r}_k}$$

Demostración: Se sigue por la representación de la diagonalización estándar

$$\mathbf{T} + \mathbf{s}\boldsymbol{\alpha} = \sum_{k=1}^p \lambda_k \frac{\mathbf{r}_k \mathbf{l}_k}{\mathbf{l}_k \cdot \mathbf{r}_k}, \quad e^{\mathbf{T} + \mathbf{s}\boldsymbol{\alpha}} = \sum_{k=1}^p e^{\lambda_k} \frac{\mathbf{r}_k \mathbf{l}_k}{\mathbf{l}_k \cdot \mathbf{r}_k}$$

■

Proposición 4.4.1.4: Supongamos las condiciones de la Proposición 4.4.1.3 y tomamos $\lambda_1 = 0$. Entonces la función de renovación y la varianza están dadas respectivamente por:

$$U(t) = 1 + c_1 t + \sum_{j=2}^p \frac{c_j}{\lambda_j} (e^{\lambda_j t} - 1)$$

$$\begin{aligned} \text{Var}(N_t) = & c_1 \left(2 - \sum_{i=2}^p \frac{c_i}{\lambda_i} \right) t + \sum_{i=2}^p \left(\frac{c_i}{\lambda_i} + 4c_1 \frac{c_i}{\lambda_i^2} - 2 \sum_{i=1}^p \frac{c_i}{\lambda_i} \right) (e^{\lambda_i t} - 1) + \sum_{i=2}^p \left(\frac{2c_i^2}{\lambda_i} - 2c_1 \frac{c_i}{\lambda_i} \right) t e^{\lambda_i t} + \\ & 2 \sum_{i \neq j, i, j > 1} \frac{c_i c_j}{\lambda_i (\lambda_j - \lambda_i)} (e^{\lambda_j t - \lambda_i t}) - \left(\sum_{i=2}^p \frac{c_i}{\lambda_i} (e^{\lambda_i t} - 1) \right)^2 \end{aligned}$$

Demostración: El resultado acerca de la función de renovación se sigue por integración simple, mientras que la fórmula para la varianza se sigue de la fórmula

$$N_t^2 = 2 \int_0^t U(t-x)u(x)dx + U(t).$$

■

En este capítulo se mostraron las distribuciones matriz exponenciales, sus propiedades, características básicas y la función de distribución de las distribuciones matriz exponenciales. También se muestran los procesos de renovación donde la distribución entre arribos \mathbf{B} es matriz exponencial y el método que permite escribir $\mathbf{T} + \mathbf{s}\boldsymbol{\alpha}$ de forma diagonal y directamente obtener la función matriz-exponencial necesitada para la densidad de renovación.

En el siguiente capítulo se dará a conocer la interpretación física que Bladt y Neuts (2003) ofrecen acerca de las distribuciones matriz exponenciales; así como también los principales resultados y propiedades que se derivan de esta interpretación.

Capítulo 5: Interpretación Física de las Distribuciones Matriz Exponenciales

5.1 Introducción

Cualquier distribución matriz exponencial tiene una representación de la forma $(\alpha, \mathbf{S}, -\mathbf{S}\mathbf{e})$. También es verdad; aunque no está explícitamente indicado que tal representación tiene $\alpha\mathbf{e} = 1$. Esta representación es usada por Bladt y Neuts (2003) en su interpretación física de las distribuciones matriz exponenciales.

Consideremos un sistema de m contenedores cada uno con una marca en cero que pueden tener una cantidad positiva, negativa o nula de líquido. Se permite que el líquido fluya determinísticamente de contenedor a contenedor. Para $i = 1, 2, \dots, m$, denotemos como $\alpha_i \in \mathbb{R}$ a la cantidad inicial de líquido en el contenedor i . Tenemos un contenedor más, el contenedor $m+1$ el cual inicialmente contiene una cantidad de líquido α_{m+1} que satisface $0 \leq \alpha_{m+1} < 1$. Supongamos que $\alpha_1 + \dots + \alpha_m + \alpha_{m+1} = 1$.

El líquido fluye del contenedor i al contenedor j , con $1 \leq i, j \leq m, i \neq j$, a una tasa constante $S_{ij} \in \mathbb{R}$. Esto significa que si el contenedor i alcanza la cantidad $v_i(t)$ al tiempo t entonces una cantidad $v_i(t)S_{ij}dt$ fluye del contenedor i al contenedor j durante $[t, t+dt)$ Para $i = 1, 2, \dots, m$, el líquido fluye al contenedor $m+1$ con una tasa constante $s_i \in \mathbb{R}$. Definimos, para $i = 1, 2, \dots, m$,

$$S_{ii} = -\sum_{\substack{j=1 \\ j \neq i}}^m S_{ij} - s_i. \quad (5.1.1)$$

Definimos $\mathbf{S} = (S_{ij})_{i,j=1,\dots,m}$. Sean los componentes del vector $\mathbf{v}(t) = (v_1(t), \dots, v_m(t))$ la cantidad de líquido en los primeros m contenedores. Los componentes de $\mathbf{v}(t)$ son funciones deterministas de t . Sea $\alpha = (\alpha_1, \dots, \alpha_m)$ Se observa que s_i es el i -ésimo componente de $-\mathbf{S}\mathbf{e}$.

Proposición 5.1.1: Para $t \geq 0$, el vector $\mathbf{v}(t)$ satisface que

$$\mathbf{v}'(t) = \mathbf{v}(t)\mathbf{S}, \quad \mathbf{v}(0) = \alpha.$$

Por lo tanto $\mathbf{v}(t) = \alpha e^{\mathbf{S}t}$.

Demostración: Obtenemos para $1 \leq i \leq m$,

$$v_i(t+dt) = v_i(t) \left[1 - \sum_{\substack{h=1 \\ h \neq i}}^m S_{ih} dt - s_i dt \right] + \sum_{\substack{j=1 \\ j \neq i}}^m v_j(t) S_{ji} dt.$$

La primera ecuación de la proposición se sigue de $S_{ii} = -\sum_{\substack{j=1 \\ j \neq i}}^m S_{ij} - s_i$ haciendo tender dt a cero;

mientras que la segunda es obvia. ■

Bladt y Neuts (2003) definieron un flujo válido con parámetros (α, \mathbf{S}) como un flujo donde la cantidad de líquido en el contenedor $m+1$ no decrece a través del tiempo y donde eventualmente todo el líquido fluye dentro del contenedor.

Definición: Un flujo (α, \mathbf{S}) es válido si y sólo si la cantidad de líquido $v_{m+1}(t)$ en el contenedor $m+1$ al tiempo t , es una función no decreciente de t y $\lim_{t \rightarrow \infty} v_{m+1}(t) = 1$.

Los flujos válidos existen. La representación (α, \mathbf{S}) de una distribución matriz exponencial determina un flujo válido. En particular, si α es no negativa y \mathbf{S} es un subgenerador, entonces (α, \mathbf{S}) es un flujo válido. Es decir, los parámetros de cualquier flujo válido corresponden a una distribución matriz exponencial.

El resultado más importante que probaron fue que cualquier variable aleatoria matriz exponencial, con representación (α, \mathbf{S}) con $\alpha \mathbf{e} = 1$, se distribuye como el tiempo de paro aleatorio de un flujo válido con parámetros (α, \mathbf{S}) . Esto es, el tiempo que toma a $v_{m+1}(t)$ alcanzar un nivel U donde U es una variable aleatoria uniforme en $[0,1]$, se distribuye como una distribución matriz exponencial con representación (α, \mathbf{S}) con $\alpha \mathbf{e} = 1$.

Teorema 5.1.1: Sea Y una variable aleatoria con distribución matriz exponencial (α, \mathbf{S}) . Consideremos el flujo válido generado por (α, \mathbf{S}) . Sea $x \in [0,1]$ y $T(x)$ el tiempo desde que el líquido comienza a fluir hasta que el contenedor $m+1$ alcanza el nivel x . Entonces Y se distribuye como $T(U)$ donde U se distribuye como una variable aleatoria uniforme $[0,1]$.

Demostración: El flujo que recibe el contenedor $m+1$ durante el intervalo $[t, t+dt)$ es la cantidad de líquido en los primeros m contenedores en el instante t por la tasa a la cual el líquido fluye al contenedor $m+1$ de cada contenedor por el intervalo de tiempo dt .

$$I_t dt = \sum_{k=1}^m v_k(t) s_k dt$$

donde U se distribuye como una variable aleatoria uniforme $[0,1]$; la probabilidad de que U esté contenida en el intervalo $[v_{m+1}(t), v_{m+1}(t) + I_t dt)$ es simplemente $I_t dt$. Por lo tanto la densidad g de $T(U)$ está dada por

$$g(t) dt = I_t dt = \sum_{k=1}^m v_k(t) s_k dt = \alpha \mathbf{e}^{\mathbf{S}t} s dt$$

Por lo tanto $T(U)$ se distribuye como Y .

■

Proposición 5.1.2: La función de distribución de una variable aleatoria F que se distribuye matriz exponencial (α, \mathbf{S}) está dada por.

$$F(x) = 1 - \alpha e^{\mathbf{S}x} e$$

donde e es un vector columna de unos.

Demostración: Para el flujo válido correspondiente a la distribución matriz exponencial (α, \mathbf{S}) los contenedores $1, \dots, m$ contendrán cada uno $\alpha e^{\mathbf{S}x}$ al tiempo x . El contenido total de los contenedores $1, \dots, m$ al tiempo x es por lo tanto $\alpha e^{\mathbf{S}x} e$. Si el contenido de todos los contenedores suma 1, el contenido en el contenedor $m+1$ al tiempo x es entonces $1 - \alpha e^{\mathbf{S}x} e$. Por lo tanto $T(U) \leq x$ si y sólo si $U \leq 1 - \alpha e^{\mathbf{S}x} e$, y

$$F(x) = \mathbb{P}(T(U) \leq x) = \mathbb{P}(U \leq 1 - \alpha e^{\mathbf{S}x} e) = 1 - \alpha e^{\mathbf{S}x} e$$

■

5.2 Cálculo con Flujos: Teoría de Renovación

En la teoría de las distribuciones tipo fase, se construye la cadena de Markov con generador infinitesimal $\mathbf{Q}^* = \mathbf{S} + \mathbf{s}\alpha$. Después de eliminar los estados superfluos, este generador puede ser siempre irreducible. El generador $\mathbf{Q}^* = \mathbf{S} + \mathbf{s}\alpha$ es útil para derivar resultados teóricos de renovación con argumentos simples. Construimos un proceso aleatorio determinista por pedazos, llamado flujo con reinicio que juega un rol análogo en la teoría de renovación con distribuciones matriz exponenciales.

Sean U_1, U_2, \dots una sucesión de variables aleatorias independientes idénticamente distribuidas con distribución uniforme en el intervalo $[0, 1]$. Sea $X_n = T(U_n)$ el tiempo correspondiente de flujo del flujo válido correspondiente a alguna distribución matriz exponencial con parámetros (α, \mathbf{S}) donde $T(\cdot)$ es definido en el Teorema 5.1.1. Entonces el proceso de renovación con tiempos de arribo matriz exponencial puede ser modelado uniendo la secuencia de los flujos terminales. Esto genera un proceso de Markov determinista por pedazos, donde los saltos corresponden a arribos ocurridos cuando el líquido del sistema se rellena a la cantidad inicial α .

Consideremos un proceso de renovación representado por un flujo con reinicio. Sea $U_i(t)$ el contenido del contenedor i al tiempo t , y definimos $z_i(t) = \mathbb{E}(U_i(t))$. Los m vectores renglón $\mathbf{U}(t)$ y $\mathbf{z}(t)$ tienen componentes $U_i(t)$ y $z_i(t)$ respectivamente.

Lema 5.2.1: La densidad de renovación $\phi(t)$ puede ser expresada como sigue:

$$\phi(t) = z(t)\mathbf{s}.$$

Demostración:

$$\phi(t)dt = \mathbb{P}(\text{arribo en } [t, t + dt])$$

$$\phi(t)dt = \mathbb{E}\left(\mathbb{P}(\text{arribo en } [t, t + dt] | U(t))\right)$$

$$\phi(t)dt = \mathbb{E}\left(\sum_k U_k(t) s_k dt\right)$$

$$\phi(t)dt = \sum_k z_k(t) s_k dt.$$

Lema 5.2.2: Para $t \geq 0$, el vector $z(t)$ está dado por

$$z(t) = \alpha \exp[(\mathbf{S} + \mathbf{s}\alpha)t].$$

Demostración 1: En el evento donde el flujo original que inició al tiempo 0 hasta el tiempo t , el contenido esperado del vector es $\alpha \exp(\mathbf{S}t)$. En el evento complementario donde ha habido un flujo que reinicia en $[0, t)$, condicionando en el tiempo u del primer reinicio, la contribución del vector esperanza $z(t)$ está dado por:

$$\int_0^t \alpha \exp[\mathbf{S}(t-u)] \mathbf{s}z(u) du.$$

Por lo tanto,

$$z(t) = \alpha \exp(\mathbf{S}t) + \int_0^t \alpha \exp[\mathbf{S}(t-u)] \mathbf{s}z(u) du.$$

Consideramos la matriz $\mathbf{V}(t)$ definida por

$$\mathbf{V}(t) = \exp(\mathbf{S}t) + \int_0^t \exp[\mathbf{S}(t-u)] \mathbf{s}z(u) du.$$

Multiplicando por $\exp(-\mathbf{S}t)$ y diferenciando nos conduce al sistema lineal de ecuaciones diferenciales

$$\mathbf{V}'(t) = \mathbf{S}\mathbf{V}(t) + \mathbf{s}z(t), \text{ con } \mathbf{V}(0) = \alpha$$

Reemplazando $z(t)$ por $\alpha\mathbf{V}(t)$, e integrando obtenemos el resultado. ■

Demostración 2: Fijamos un tiempo t . En el evento A donde no hay arribos en $[t, t + dt)$, tenemos que

$$\mathbb{E}(U_k(t + dt) 1_A | \mathbf{U}(t)) = U_k(t) + \sum_{j \neq k} U_j(t) S_{jk} dt - \sum_{j \neq k} U_k(t) S_{kj} dt - U_k(t) s_k dt$$

cuando el desarrollo de $U_k(t)$ en A es determinista y el contenido del contenedor k al tiempo $t + dt$ es el contenido del tiempo t más el flujo neto. Tomando la esperanza,

$$\mathbb{E}(U_k(t+dt)1_A) = z_k(t) + \sum_{j \neq k} z_j(t)S_{jk}dt - \sum_{j \neq k} z_k(t)S_{kj}dt - z_k(t)s_k dt$$

En el evento complementario, A^c , habrá un salto en algún punto en $[t, t + dt)$. Aquí sabemos que la esperanza condicional $\mathbb{E}(U_k(t+dt)|A^c) = \alpha_k + o(dt)$. Así

$$z_k(t+dt) = \mathbb{E}(1_A U_k(t+dt) + 1_{A^c} U_k(t+dt))$$

$$z_k(t+dt) = \mathbb{E}(1_A U_k(t+dt)|U(t)) + \mathbb{E}(U_k(t+dt)|A^c)P(A^c)$$

$$z_k(t+dt) = z_k(t) + \sum_{j \neq k} z_j(t)S_{jk}dt - \sum_{j \neq k} z_k(t)S_{kj}dt - z_k(t)s_k dt + \alpha_k \sum_j z_j(t)s_j dt$$

Entonces, si $S_{ii} = -\sum_{i \neq j} S_{ij} - s_i$, tenemos que

$$z_k(t+dt) = z_k(t) + \sum_j z_j(t)S_{jk}dt + \sum_i z_i(t)s_i dt \alpha_k.$$

Esto nos lleva a las ecuaciones diferenciales

$$z'_k(t) = \sum_j (S_{jk} + s_j \alpha_k) z_j(t),$$

para $1 \leq k \leq m$, cuya solución es

$$z(t) = \mathbf{a} \exp[(\mathbf{S} + \mathbf{s}\mathbf{a})t].$$

■

Teorema 5.2.1: La densidad de renovación $\phi(t)$ correspondiente a $F(\cdot)$ está dada por

$$\phi(t) = \mathbf{a} \exp[(\mathbf{S} + \mathbf{s}\mathbf{a})t] \mathbf{s}.$$

Demostración: Se obtiene al combinar los resultados de los dos lemas anteriores.

■

En cuanto a distribuciones tipo fase podemos explotar la interpretación de flujos para obtener propiedades adicionales para procesos de renovación con tiempos entre arribos matriz exponenciales. Así fácilmente se obtiene el siguiente teorema.

Teorema 5.2.2: El tiempo de vida residual al tiempo x , V_x , el cual es definido como el tiempo desde x hasta el próximo arribo, tiene una distribución matriz exponencial con representación

$$(\alpha \exp((\mathbf{S} + s)x), \mathbf{S})$$

Demostración Si el contenido de los m contenedores es $U(x)$ al tiempo x y el próximo arribo ocurre en $[x+y, x+y+dy)$, entonces del tiempo x al $x+y$ el líquido fluye de manera determinista, entonces el contenido de los contenedores al tiempo $x+y$ es $U(x) \exp(\mathbf{S}y)$. Así la probabilidad de renovación en $[x+y, x+y+dy)$ dado $U(x)$, es $U(x) \exp(\mathbf{S}y) \mathbf{s}^0 dy$. Por lo tanto

$$\mathbb{P}(V_x \in [y, y+dy)) = \mathbb{E}(\mathbb{P}(V_x \in [y, y+dy) | U(x))) = \mathbb{E}(U(x) \exp(\mathbf{S}y) \mathbf{s}^0 dy)$$

$$\mathbb{P}(V_x \in [y, y+dy)) = z(x) \exp(\mathbf{S}y) \mathbf{s}^0 dy.$$

■

5.3 Modelo de Teoría del Riesgo

A continuación daremos una pequeña descripción del modelo de Asmussen y Bladt (1999).

La reserva de riesgo al tiempo t de un fondo de seguro es R_t , con $R_0 = u > 0$. La reserva crece a una tasa $p(r)$ la cual depende de la reserva actual r , y es equivalente a que, entre las reclamaciones, la reserva se desarrolla determinísticamente como $\left(\frac{d}{dt}\right) R_t = p(R_t)$. La reserva decrece por saltos iguales al tamaño de las reclamaciones. Suponemos que el tamaño de las reclamaciones sucesivas es independiente con la misma distribución común $B(\cdot)$.

Supongamos que $B(\cdot)$ es una distribución matriz exponencial con parámetros (\mathbf{u}, \mathbf{S}) y es representada por el correspondiente flujo válido. Supongamos que las reclamaciones llegan conforme a un proceso Poisson homogéneo con parámetro β . El monto de las reclamaciones corresponde a la duración del flujo. Correspondiendo a cada punto en el salto hacia abajo de la reclamación, hay un vector de contenidos bien definido en el modelo de flujo. Denotaremos por $v_i(u)$ a la cantidad esperada de líquido en el contenedor i en el primer momento que descienda del nivel u , condicional en la reserva al tiempo 0 siendo u .

Teorema 5.3.1: Para $u > 0$ y $1 \leq i \leq m$, las cantidades $v_i(u)$ satisfacen el sistema de ecuaciones diferenciales no lineales

$$-p(u)v_i'(u) = \beta \alpha_i + v_i(u) \left(\sum_{j=1}^m p(u)v_j(u)s_j - \beta \right) + \sum_{j=1}^m p(u)v_j(u)S_{ji} \quad (5.3.1)$$

Demostración: Tomaremos en cuenta todo lo que puede ocurrir en $[0, dt)$. Con probabilidad βdt , habrá una reclamación en este intervalo. En este caso u inmediatamente desciende por debajo del nivel de u y el contenido de i a ese tiempo es α_i . Con la probabilidad del evento

complementario $1 - \beta dt$, que no haya algún arribo en $[0, dt)$ y la reserva crece a $u + p(u)dt$. Para que exista un descenso por debajo del nivel de u , debe existir un descenso por debajo del nivel de $u + p(u)dt$. Consideremos la reclamación que resulta en el primer descenso por debajo de $u + p(u)dt$. Existen dos alternativas: que el modelo del flujo para la reclamación termine en $[u, u + p(u)dt)$ o que continúe y también descienda por debajo del nivel u .

Consideremos el caso en el cual el modelo del flujo para la reclamación termine en $[u, u + p(u)dt)$. Sea U_k^1 el contenido del contenedor k sobre el nivel $u + p(u)dt$ y sea $\mathbf{U}^1 = (U_1^1, \dots, U_m^1)$. Entonces la probabilidad de que ocurra este caso es la probabilidad de que el descenso se detenga en $(u, u + p(u)dt]$

$$\mathbb{E}\left(\sum_k U_k^1 s_k p(u)dt\right) = \sum_k v_k(u + p(u)dt) s_k p(u)dt.$$

Condicionando en este evento, el contenido esperado del contenedor i hasta que descienda del nivel u es simplemente $v_i(u)$ y el proceso esencialmente inició sobre el nivel u . En el otro caso calculamos el evento A que el proceso cruce por debajo del nivel u . Condicionando en el contenido \mathbf{U}^1 del nivel $u + p(u)dt$ obtenemos que en A el contenido cuando cruce por debajo del nivel u del contenedor i , U_i^2 , debe satisfacer

$$\mathbb{E}\left(U_i^2 1_A \mid \mathbf{U}^1\right) = U_i^1 + \sum_{j \neq i} U_j^1 S_{ji} p(u)dt - \sum_{j \neq i} U_i^1 S_{ij} p(u)dt - U_i^1 p(u)dt s_i.$$

Tomando esperanzas obtenemos

$$\mathbb{E}\left(\mathbb{E}\left(U_i^2 1_A \mid \mathbf{U}^1\right)\right) = \mathbb{E}\left(U_i^1\right) + \mathbb{E}\left(\sum_{j \neq i} U_j^1 S_{ji} p(u)dt\right) - \mathbb{E}\left(\sum_{j \neq i} U_i^1 S_{ij} p(u)dt\right) - \mathbb{E}\left(U_i^1 p(u)dt s_i\right)$$

$$\mathbb{E}\left(U_i^2 1_A\right) = \mathbb{E}\left(U_i^1\right) + \sum_{j \neq i} \mathbb{E}\left(U_j^1\right) S_{ji} p(u)dt - \sum_{j \neq i} \mathbb{E}\left(U_i^1\right) S_{ij} p(u)dt - \mathbb{E}\left(U_i^1\right) p(u)dt s_i$$

$$\mathbb{E}\left(U_i^2 1_A\right) = v_i(u + p(u)dt) + \sum_{j \neq i} v_j(u + p(u)dt) S_{ji} p(u)dt - \sum_{j \neq i} v_i(u + p(u)dt) S_{ij} p(u)dt - v_i(u + p(u)dt) p(u)dt s_i$$

$$\mathbb{E}\left(U_i^2 1_A\right) = v_i(u + p(u)dt)(1 + S_{ii} p(u)dt) + \sum_{j \neq i} v_j(u + p(u)dt) S_{ji} p(u)dt$$

$$\mathbb{E}\left(U_i^2 1_A\right) = v_i(u + p(u)dt) + \sum_{j=1}^m v_j(u + p(u)dt) S_{ji} p(u)dt.$$

Ahora agregando las contribuciones al contenido esperado de i , tenemos que

$$v_i(u) = \alpha_i \beta dt + (1 - \beta dt) \left(\sum_{j=1}^m v_j(u + p(u)dt) (\delta_{ji} + p(u)dt S_{ji}) + \sum_{j=1}^m v_j(u + p(u)dt) s_j \cdot p(u)dt \cdot v_i(u) \right) \quad (5.3.2)$$

para $1 \leq i \leq m$. Ahora usando $v_j(u + p(u)dt) = v_j(u) + v'_j(u)p(u)dt$ y simplificando, obtenemos (5.3.1).

Sea ahora $\psi(u)$ la probabilidad condicional que, comenzando en u , la reserva eventualmente cruce el nivel 0. En términos de las funciones $v_i(u)$, la probabilidad de ruina $\psi(u)$ puede ser obtenida de la siguiente manera:

Teorema 5.3.2: La probabilidad de ruina está dada por $\psi(u) = \sum_{i=1}^m \lambda_i(u)$, donde las funciones $\lambda_j(u)$ son soluciones del sistema de ecuaciones diferenciales

$$\lambda'_i(t) = \sum_{j=1}^m \lambda_j(t) v_i(u-t) s_j + \sum_{j=1}^m \lambda_j(t) S_{ji} \quad (5.3.3)$$

para $1 \leq i \leq m$ y sujeto a $\lambda_i(0) = v_i(u)$

Demostración:

Comenzando en el nivel u , consideramos la primera reclamación cuando el proceso llega por primera vez por debajo del nivel u . Existen dos alternativas. El primer toma toda la trayectoria del proceso por debajo de 0 o la reserva después de la reclamación es $u - u' > 0$, donde u' es el monto causado por la reclamación. En el último de los casos, repetimos el procedimiento iniciando con reserva $u - u'$ y posiblemente obtenemos un segundo punto $u - u'' = u - u' + u' - u''$ donde u'' es el monto de la primera reclamación la cual cruzó por debajo del nivel $u - u'$. Continuando de esta manera descomponemos la trayectoria que nos conduce a la ruina de acuerdo con los valores sucesivos de u', u'', \dots . Nos referimos a la variable independiente generada por u', u'', \dots como el tiempo virtual para distinguirlo del tiempo real del proceso de riesgo.

Juntamos los flujos correspondientes a los sucesivos trayectorias por debajo de u . En el modelo de riesgo con reserva inicial u , la ruina ocurre si y sólo si esta concatenación de trayectorias por debajo de u dura al menos hasta el tiempo virtual u .

Para $t \geq 0$, sea $U_j(t)$ el contenido del contenedor al tiempo virtual t y $\lambda_j(t)$ la cantidad esperada del contenedor j . Sea $\mathbf{U}(t) = (U_1(t), \dots, U_m(t))$. Sea A el evento en el que el flujo actual al tiempo virtual t continua hasta $[t, t + dt)$. Entonces

$$\mathbb{E}(U_i(t + dt) | \mathbf{U}(t)) = \mathbb{E}(U_i(t + dt) 1_A | \mathbf{U}(t)) + \mathbb{E}(U_i | \mathbf{U}(t), A^c) \mathbb{P}(A^c | \mathbf{U}(t)) \quad (5.3.4)$$

En A y condicionando sobre $\mathbf{U}(t)$ obtenemos que

$$\mathbb{E}\left(U_i(t+dt)1_A \mid \mathbf{U}(t)\right) = U_i(t) + \sum_{j \neq i} U_j(t)S_{ji}dt - \sum_{j \neq i} U_i(t)S_{ij}dt - U_i(t)s_i dt$$

$$\mathbb{E}\left(U_i(t+dt)1_A \mid \mathbf{U}(t)\right) = \sum_{j=1}^m (\delta_{ij} + S_{ji}dt)U_j(t).$$

Para el evento complementario,

$$\mathbb{P}\left(A^c \mid \mathbf{U}(t)\right) = \sum_j U_j(t)s_j dt$$

mientras que

$$\mathbb{E}\left(U_i(t+dt) \mid A^c, \mathbf{U}(t)\right) = v_i(u-t)$$

entonces el tiempo virtual t corresponde al nivel $u-t$ de la reserva.

Introduciendo estas expresiones en (5.3.4) y tomando esperanzas obtenemos que

$$\lambda_i(t+dt) = \sum_{j=1}^m \lambda_j(t) [\delta_{ji} + S_{ji}dt + s_j dt \cdot v_i(u-t)], \text{ para } 1 \leq i \leq m. \quad (5.3.5)$$

Para flujos válidos, la suma de los contenidos al tiempo virtual t de los contenedores i , $1 \leq i \leq m$, es precisamente la probabilidad que el flujo no haya terminado. Por lo tanto, $\sum_{j=1}^m \lambda_j(u)$ es la probabilidad que al tiempo virtual u un flujo este activo. Esta es precisamente la probabilidad de ruina. ■

5.4 Procesos de Arribo Racionales

Assmussen y Bladt (1997) desarrollaron una interpretación física para las distribuciones matriz exponenciales para los procesos de arribo racionales. Los procesos de arribo racionales son una extensión de los procesos de arribo Markovianos. Un proceso de arribo Markoviano es un proceso puntual Markoviano definido en un espacio de estados finito donde la transición entre estados sin un arribo o evento ocurrido es gobernada por una matriz \mathbf{D} . Si $\boldsymbol{\alpha}$ es la distribución inicial de los estados entonces $(\boldsymbol{\alpha}, \mathbf{C}, \mathbf{D})$ es una representación del proceso de arribo Markoviano. Los procesos de arribo racionales son procesos puntuales que permiten tiempos entre arribos distribuidos matriz exponenciales.

Consideremos lo siguiente. Sea una variable aleatoria X definida en los números reales no negativos tiene función de distribución F . Para $t \geq 0$, sea $F_t(x) = F(x-t) - F(x)$ la distribución del tiempo de vida residual $X-t$. Si $span(F)$ denota el vector espacio de medida, consistente en todas las combinaciones lineales de F_t , donde $t \geq 0$, entonces tenemos que una distribución F es una distribución matriz exponencial si y solo si $span(F)$ es finito dimensional.

La dimensionalidad finita permite representar un proceso de arribo racional como un proceso de Markov determinista por pedazos de dimensión p $\mathbf{A}(t) = (A_1(t), A_2(t), \dots, A_p(t))$, donde $t \geq 0$. Un proceso de Markov determinista evoluciona determinísticamente conforme a una ecuación diferencial multidimensional y a un tiempo aleatorio los puntos cambian de estado. El espacio de estados es el conjunto de todos los posibles vectores $\mathbf{A}(t)$ donde $t \geq 0$. Entre los saltos el proceso de Markov determinista por pedazos que representa el proceso de arribo racional evoluciona conforme a la ecuación diferencial

$$\frac{d}{dt}(\mathbf{a}(t)) = \mathbf{a}(t)\mathbf{C} - (\mathbf{a}(t)\mathbf{C}\mathbf{e})\mathbf{a}(t) \quad (5.4.1)$$

donde $\mathbf{a}(t) = (a_1(t), a_2(t), \dots, a_p(t))$ es el estado del proceso al tiempo $t \geq 0$ con $\mathbf{a}(t)\mathbf{e} = 1$. La solución de la ecuación diferencial (5.4.1), la cual está definida para $t \geq 0$, está dada por

$$\mathbf{a}(t) = \frac{\mathbf{a}(0)\exp(\mathbf{C}t)}{\mathbf{a}(0)\exp(\mathbf{C}t)\mathbf{e}}$$

Cuando el proceso de Markov determinista por pedazos está en el estado $\mathbf{a}(t)$, un salto aleatorio ocurre con intensidad $\mathbf{a}(t)\mathbf{D}$. Si el salto ocurre cuando $t = t^*$ el proceso evoluciona determinísticamente conforme a (5.4.1) entonces inicia de nuevo con

$$\frac{\mathbf{a}(t^*)\mathbf{D}}{\mathbf{a}(t^*)\mathbf{D}\mathbf{e}}$$

como el estado inicial.

Asmussen y Bladt (1999) representan la evolución de un proceso de arribos racional con una representación de orbitas, esto es, como una trayectoria en el subespacio $(p-1)$ dimensional $x_1 + x_2 + \dots + x_p = 1$ de \mathbb{R}^p gobernado por la ecuación paramétrica $\mathbf{x}(t) = \mathbf{A}(t)$. La trayectoria entera $\mathbf{x}(t)$, para $t \geq 0$, es una sucesión de orbitas en las cuales el vector $\mathbf{x}(t)$ satisface (5.4.1). Cuando un salto ocurre la trayectoria cambia a otra órbita. Si el proceso de arribos racional es un proceso de renovación la trayectoria reiniciará en la misma órbita en el mismo punto inicial después de cada salto.

En los procesos de arribos racionales si el estado del proceso al tiempo t , esto es, $\mathbf{a}(t)$ es conocido, la evolución futura es independiente del pasado.

A continuación daremos una interpretación física de los procesos de arribo racionales. Sean \mathbf{C} y \mathbf{D} las matrices de coeficientes de un proceso de arribo racional junto con el vector de contenido inicial. Entonces, con $\mathbf{s} = \mathbf{D}\mathbf{e}$, la densidad conjunta g_n para los primeros n tiempos entre arribos es

$$g_n(x_1, \dots, x_n) = \boldsymbol{\alpha} \exp(\mathbf{C}x_1)\mathbf{D} \exp(\mathbf{C}x_2)\mathbf{D} \dots \mathbf{D} \exp(\mathbf{C}x_n)\mathbf{s}.$$

Ahora consideremos el siguiente esquema de flujo. Inicialmente, los contenedores $1, \dots, m$ contienen niveles $\alpha_1, \dots, \alpha_m$. La tasa de flujo entre los contenedores son los elementos de \mathbf{C} . Así también existe un flujo creciente que entra al contenedor $m+1$. El flujo del contenedor i al contenedor $m+1$ es $s_i = (-\mathbf{C}\mathbf{e})_i = (\mathbf{D}\mathbf{e})_i$. Cuando un flujo se detiene debido a un arribo y el vector contenido de contenedores $1, \dots, m$ justo antes el arribo es $\boldsymbol{\beta}$, entonces los contenedores son rellenados al vector contenido $\frac{\boldsymbol{\beta}\mathbf{D}}{(\boldsymbol{\beta}\mathbf{D}\mathbf{e})}$. Este esquema de flujo nos provee una interpretación física de los procesos de arribo racionales. Consideremos los flujos como fueron descritos. Entonces,

$$g_2(x_1, x_2) = g_2(x_2|x_1)g_2(x_1) = \frac{\boldsymbol{\alpha}\exp(\mathbf{C}x_1)\mathbf{D}}{\boldsymbol{\alpha}\exp(\mathbf{C}x_1)\mathbf{D}\mathbf{e}}\exp(\mathbf{C}x_2)\mathbf{s} \cdot \boldsymbol{\alpha}\exp(\mathbf{C}x_1)\mathbf{s}$$

$$g_2(x_1, x_2) = \boldsymbol{\alpha}\exp(\mathbf{C}x_1)\mathbf{D}\exp(\mathbf{C}x_2)\mathbf{s}$$

Por inducción, la densidad conjunta de los primeros n tiempos entre arribos en el esquema de flujo g_n es

$$g_n(x_1, \dots, x_n) = \boldsymbol{\alpha}\exp(\mathbf{C}x_1)\mathbf{D}\dots\mathbf{D}\exp(\mathbf{C}x_n)\mathbf{s}$$

■

En este capítulo se mostró la interpretación física que Bladt y Neuts (2003) ofrecen acerca de las distribuciones matriz exponenciales definiendo el flujo válido para poder definir la relación entre Y que es una variable aleatoria con distribución matriz exponencial $(\boldsymbol{\alpha}, \mathbf{S})$ y $T(U)$ donde U se distribuye como una variable aleatoria uniforme $[0,1]$. También se mostraron resultados importantes sobre los procesos de renovación con tiempos de arribo matriz-exponencial modelados como secuencias de flujos terminales

También se mostró el modelo de riesgo que Asmussen y Bladt (1999) modelaron bajo la interpretación física de flujos y la interpretación física para los procesos de arribo racionales con una representación de orbitas.

En el siguiente capítulo se mostrarán las distribuciones matriz exponenciales multivariadas como una extensión natural del caso univariado; así como también la teoría de renovación.

Capítulo 6: Distribuciones Matriz Exponenciales Multivariadas

6.1 Fracciones Continuas

Una fracción continua es una expresión de la forma

$$d_0 + \frac{c_1}{d_1 + \frac{c_2}{d_2 + \frac{c_3}{d_3 + \dots}}}$$

Es conveniente usar una notación más compacta

$$d_0 + \frac{c_1}{|d_1} + \frac{c_2}{|d_2} + \frac{c_3}{|d_3} + \dots$$

Se dice que una fracción continua es finita si la suma de arriba contiene un número finito de términos. De interés particular para nuestro análisis son las fracciones C -continuas, las cuales tienen expresiones de la forma

$$1 + \frac{c_1 s^{r_1}}{|1} + \frac{c_2 s^{r_2}}{|1} + \frac{c_3 s^{r_3}}{|1} + \dots$$

Si consideramos la función generadora de momentos $M(s)$ de una variable aleatoria distribuida matriz exponencial, entonces su expansión en serie de potencias es

$$M(s) = 1 + \mu_1 s + \mu_2 s^2 + \dots$$

donde μ_i esta definida como $\mu_i = \frac{M_i}{i!}$, $i = 0, 1, \dots$

Notemos que $\mu_i > 0$ para toda i . De acuerdo con Perron (1957), cualquier serie de potencias con término constante 1 corresponde únicamente a una fracción C -continua. Si además la serie es una expansión en series de potencias de una función racional, entonces la correspondiente fracción continua es finita. Por lo tanto son tratables las fracciones C -continuas, donde $r_i = 1$ para toda i

Lema 6.1.1: Sea $M(s)$ la función generadora de momentos de una variable aleatoria distribuida matriz exponencial. Entonces la expansión en series de potencias de $M(s) = 1 + \mu_1 s + \mu_2 s^2 + \dots$ corresponde únicamente a una fracción C -continua.

Demostración: Sabemos por Perron (1957) que existe una fracción C -continua finita, entonces lo que tenemos que demostrar es que $r_i = 1$ para toda i .

Por lo tanto

$$B_0(s) = 1 + \mu_1 s + \mu_2 s^2 + \dots = 1 + \frac{c_1 s^{r_1}}{B_1(s)},$$

donde

$$B_1(s) = 1 + \frac{c_2 s^{r_2}}{1} + \frac{c_3 s^{r_3}}{1} + \dots + \frac{c_v s^{r_v}}{1}.$$

El primer término de la expansión en series de potencias $B_0(s) - 1$ es $\mu_1 s$. Por lo tanto $r_i = 1$. Así

$$B_1(s) = \frac{\mu_1(s)}{B_0(s) - 1} = \frac{1}{1 + \mu'_1 s + \mu'_2 s^2 + \dots},$$

donde $\mu'_i = \frac{\mu_i}{\mu_1} \neq 0$ para toda i . Entonces el primer término en la expansión en series de potencias de $B_1(s) - 1$ es otra vez de primer orden. Continuando de esta forma hasta que obtenemos que $B_v(s) = 1$ probamos el resultado. ■

Para cualquier fracción continua nos podemos aproximar a ella por una de orden menor n , donde $n < v$ para el caso finito,

$$\frac{C_n}{D_n} = d_0 + \frac{c_1}{d_1} + \frac{c_2}{d_2} + \frac{c_3}{d_3} + \dots + \frac{c_n}{d_n}.$$

La fracción finita $\frac{C_n}{D_n}$ es llamada una aproximación a la fracción continua. El siguiente esquema de recursión es utilizado para las aproximaciones de diferentes órdenes

$$C_n = d_n C_{n-1} + c_n C_{n-2}$$

$$D_n = d_n D_{n-1} + c_n D_{n-2}$$

con las siguientes restricciones $C_{-1} = 1$, $C_0 = d_0$ y $D_{-1} = 0$, $D_0 = 1$. Si aplicamos estas ecuaciones recursivas a las fracciones C -continuas obtenemos que los polinomios $C_n(s)$ y $D_n(s)$ satisfacen

$$C_n(s) = 1 + \sum_{j=1}^n s^j \left(\sum_{0 \leq i_1 < i_2 < \dots < i_j < n-j} c_{i_1+1} c_{i_2+2} \dots c_{i_j+j} \right)$$

$$D_n(s) = 1 + \sum_{j=1}^n s^j \left(\sum_{1 \leq i_1 < i_2 < \dots < i_j < n-j} c_{i_1+1} c_{i_2+2} \dots c_{i_j+j} \right),$$

donde asumimos que las sumas sobre conjuntos vacíos. Así podemos escribir

$$C_{2n}(s) = 1 + \alpha_{n,1}s + \alpha_{n,2}s^2 + \dots + \alpha_{n,n}s^n$$

$$D_{2n}(s) = 1 + \beta_{n,1}s + \beta_{n,2}s^2 + \dots + \beta_{n,n}s^n$$

$$C_{2n-1}(s) = 1 + \gamma_{n,1}s + \gamma_{n,2}s^2 + \dots + \gamma_{n,n}s^n$$

$$D_{2n-1}(s) = 1 + \delta_{n,1}s + \delta_{n,2}s^2 + \dots + \delta_{n,n}s^n$$

En particular tenemos que $\beta_{n,n} = c_2 c_4 \dots c_{2n}$ y $\gamma_{n,n} = c_1 c_3 \dots c_{2n-1}$ las cuales se convierten en expresiones útiles.

Ahora consideremos la función generadora de momentos $M(s)$ y su expansión en serie de potencias en términos de los momentos reducidos. La expansión en potencias de la aproximación de k -ésimo orden es $\frac{C_k}{D_k}$. Entonces los primeros k términos de ambas series de potencias coinciden Perron (1957). Por lo tanto podemos escribir

$$\frac{C_k}{D_k} = 1 + \mu_1 s + \mu_2 s^2 + \dots + \mu_k s^k + \tilde{\mu}_{k+1} s^{k+1} + \dots,$$

donde $\tilde{\mu}_i$ son algunas constantes. Ahora insertando la expresión de arriba para $C_k(s)$ y $D_k(s)$ con $k = 2n$ y $k = 2n-1$ tenemos que

$$\frac{1 + \alpha_{n,1}s + \dots + \alpha_{n,n}s^n}{1 + \beta_{n,1}s + \dots + \beta_{n,n}s^n} = 1 + \mu_1 s + \dots + \mu_{2n} s^{2n} + \tilde{\mu}_{2n+1} s^{2n+1} + \dots \quad (6.1.1)$$

$$\frac{1 + \gamma_{n,1}s + \dots + \gamma_{n,n}s^n}{1 + \delta_{n,1}s + \dots + \delta_{n,n}s^n} = 1 + \mu_1 s + \dots + \mu_{2n-1} s^{2n-1} + \tilde{\mu}'_{2n} s^{2n} + \dots$$

Podemos ahora resolver para las constantes $\alpha_{n,i}$, $\beta_{n,i}$, $\gamma_{n,i}$ y $\delta_{n,j}$, $i = 1, \dots, n$, $j = 1, \dots, n-1$; multiplicando los numeradores de las fracciones por el lado derecho de la ecuación e igualando los coeficientes de los términos s^i , $i = n+1, n+2, \dots, 2n$.

Obtenemos el siguiente sistema de ecuaciones

$$\begin{aligned}
 0 &= \mu_{n+1} + \mu_1 \beta_{n,n} + \mu_2 \beta_{n,n-1} + \dots + \mu_n \beta_{n,1} \\
 0 &= \mu_{n+2} + \mu_2 \beta_{n,n} + \mu_3 \beta_{n,n-1} + \dots + \mu_{n+1} \beta_{n,1} \\
 &\quad \vdots \\
 0 &= \mu_{2n} + \mu_n \beta_{n,n} + \mu_{n+1} \beta_{n,n-1} + \dots + \mu_{2n-1} \beta_{n,1}
 \end{aligned}$$

que es lo mismo que

$$\begin{aligned}
 -\mu_{n+1} &= \mu_1 \beta_{n,n} + \mu_2 \beta_{n,n-1} + \dots + \mu_n \beta_{n,1} \\
 -\mu_{n+2} &= \mu_2 \beta_{n,n} + \mu_3 \beta_{n,n-1} + \dots + \mu_{n+1} \beta_{n,1} \\
 &\quad \vdots \\
 -\mu_{2n} &= \mu_n \beta_{n,n} + \mu_{n+1} \beta_{n,n-1} + \dots + \mu_{2n-1} \beta_{n,1}.
 \end{aligned}$$

Por la regla de Cramér,

$$\beta_{n,i} = \frac{\begin{vmatrix} \mu_1 & \dots & -\mu_{n+1} & \dots & \mu_n \\ \mu_2 & \dots & -\mu_{n+2} & \dots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_n & \dots & -\mu_{2n} & \dots & \mu_{2n-1} \end{vmatrix}}{\begin{vmatrix} \mu_1 & \mu_2 & \dots & \mu_n \\ \mu_2 & \mu_3 & \dots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n-1} \end{vmatrix}}.$$

En particular, para $\beta_{n,n}$ tenemos que

$$\beta_{n,n} = (-1)^{n-1} \frac{\psi_{n+1}}{\phi_n},$$

donde ψ_n y ϕ_n son los determinantes de Hankel definidos por

$$\phi_n = \begin{vmatrix} \mu_1 & \mu_2 & \dots & \mu_n \\ \mu_2 & \mu_3 & \dots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n-1} \end{vmatrix} \quad \text{y} \quad \psi_n = \begin{vmatrix} \mu_2 & \mu_3 & \dots & \mu_n \\ \mu_3 & \mu_4 & \dots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n-2} \end{vmatrix} \quad (6.1.2)$$

para toda $n = 1, 2, \dots$ para ϕ_n y para $n = 2, 3, \dots$ para ψ_n . Así

$$\beta_{n,n}\phi_n = (-1)^{n-1}\psi_{n+1}.$$

Similarmente igualando para $\frac{C_{2n-1}(s)}{D_{2n-1}(s)}$. Tenemos que

$$\gamma_{n,n}\psi_n = (-1)^{n-1}\phi_n, \quad n = 1, 2, \dots$$

con $\psi_1 = 1$.

Insertando las expresiones $\beta_{n,n} = c_2c_4 \dots c_{2n}$ y $\gamma_{n,n} = c_1c_3 \dots c_{2n-1}$ tenemos que

$$c_2c_4 \dots c_{2n}\phi_n = (-1)^{n-1}\psi_{n+1}$$

$$c_1c_3 \dots c_{2n-1}\psi_n = (-1)^{n-1}\phi_n$$

Si la función generadora de momentos es una función racional de n -ésimo orden, entonces la fracción continua correspondiente tiene al menos $2n$ términos distintos de cero $c_0 = 1$, c_1, c_2, \dots, c_{2n} y $c_v = 0$ para $v > 2n$. Entonces $\phi_{n+1} = 0$ por la segunda ecuación y por lo tanto $\psi_{n+2} = 0$ por la primera ecuación. Por lo tanto todos los determinantes de Hankel de orden mayor $\phi_v = 0$ para $v > n$ y $\psi_v = 0$ para $v > n+1$. Además, podemos reescribir las constantes c_1, c_2, \dots en términos de las matrices de Hankel definiendo $\phi_0 = 1$ y

$$c_1 = \phi_1, \quad c_{2n} = -\frac{\psi_{n+1}\phi_{n-1}}{\psi_n\phi_n}, \quad c_{2n} = -\frac{\phi_{n+1}\psi_n}{\psi_{n+1}\phi_n}$$

cuando los coeficientes sean distintos de cero.

El término $\alpha_{n,n}$ puede ser calculado similarmente considerando las mismas ecuaciones pero igualando de s^i , $i = n, n+2, \dots, 2n-1$ en lugar de s^i , $i = n+1, n+2, \dots, 2n$. Así tenemos que

$$\alpha_{n,n} = \frac{\begin{vmatrix} 1 & \mu_1 & \mu_2 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \mu_{n+2} & \cdots & \mu_{2n-1} \end{vmatrix}}{\begin{vmatrix} \mu_2 & \mu_3 & \cdots & \mu_n \\ \mu_3 & \mu_4 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n-1} \end{vmatrix}}$$

Si la función generadora de momentos es una función racional de orden n , entonces $\alpha_{n,n}$ es el coeficiente de s^n en el numerador y debe ser cero. Por lo tanto también hemos probado que el determinante de Hankel

$$H_n = \begin{vmatrix} 1 & \mu_1 & \mu_2 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \mu_{n+2} & \cdots & \mu_{2n-1} \end{vmatrix}$$

es cero cuando n es el orden de la función racional. El rango de la matriz correspondiente a H_n es $n-1$ entonces el menor subdeterminante derecho es distinto de cero de acuerdo con el análisis previo. Por lo tanto $H_{n-1} \neq 0$ y por un argumento similar podemos concluir que $H_m \neq 0$ cuando $m < n$. El orden de la distribución matriz exponencial puede ser verificando los determinantes siendo la primera vez que son distintos de cero.

Teorema 6.1.1: Consideremos una variable aleatoria X con distribución matriz exponencial y con momentos reducidos $\mu_i = \frac{\mathbb{E}(X^i)}{i!}$. Entonces la función generadora de momentos racional de X puede ser escrita como una fracción C -continua finita

$$1 + \frac{c_1 s}{1} + \frac{c_2 s^2}{1} + \cdots + \frac{c_{2n}}{1}.$$

Los coeficientes c_i pueden ser calculados en términos de los determinantes de Hankel (6.1.2) como sigue:

$$c_1 = \phi_1, \quad c_{2n} = -\frac{\psi_{n+1}\phi_{n-1}}{\psi_n\phi_n}, \quad c_{2n+1} = -\frac{\phi_{n+1}\psi_n}{\psi_{n+1}\phi_n},$$

donde $\phi_0 = 1$. Los determinantes de Hankel $\phi_m = 0$ para $m > n$ y $\psi_m = 0$ para $m > n+1$.

6.2 Distribuciones Matriz Exponenciales Multivariadas

Definiremos las distribuciones matriz exponenciales multivariadas como una extensión natural del caso univariado (Bladt y Nielsen. 2008)

Definición: Un vector aleatorio no negativo $\mathbf{X} = (X_1, \dots, X_n)$ de dimensión n tiene una distribución matriz exponencial multivariada si la transformada de Laplace conjunta $L(\mathbf{s}) = \mathbb{E}[\exp(-\langle \mathbf{s}, \mathbf{X} \rangle)]$ es una función racional multidimensional, esto es, una fracción entre dos polinomios multidimensionales. Aquí $\langle \cdot, \cdot \rangle$ denota el producto punto en \mathbb{R}^n y $\mathbf{s} = (s_1, \dots, s_n)$.

Teorema 6.2.1: Un vector $\mathbf{X} = (X_1, \dots, X_n)$ tiene una distribución matriz exponencial multivariada si y solo si $\langle \mathbf{a}, \mathbf{X} \rangle = \sum_{i=1}^n a_i X_i$ tiene una distribución matriz exponencial univariada para todos los vectores no negativos $\mathbf{a} \neq \mathbf{0}$.

Demostración: Supongamos que \mathbf{X} tiene una distribución matriz exponencial multivariada. Entonces usando

$$\mathbb{E}[\exp(-s \langle \mathbf{a}, \mathbf{X} \rangle)] = \mathbb{E}[\exp(-\langle s\mathbf{a}, \mathbf{X} \rangle)]$$

Concluimos que el lado izquierdo de la ecuación es una función racional de s entonces el lado derecho es, por definición, racional en $s\mathbf{a}$.

Ahora suponemos que $\langle \mathbf{a}, \mathbf{X} \rangle$ tiene una transformada de Laplace racional, y por lo tanto tiene función generadora de momentos, para toda $\mathbf{a} \neq \mathbf{0}$ no negativa. Las dimensiones de las representaciones para $\langle \mathbf{a}, \mathbf{X} \rangle$ están restringidas para algunas m . Suponemos lo contrario. La dimensión de cualquier distribución no será afectada normalizando \mathbf{a} por $\mathbf{a}\mathbf{e}$. Entonces \mathbf{a} es no negativa y podemos restringir la atención al conjunto compacto $\{\mathbf{a} \geq \mathbf{0} : \mathbf{a}\mathbf{e} = 1\}$. Si la dimensión no está restringida entonces existe una sucesión $\mathbf{a}_n \rightarrow \mathbf{a}_0$ en el conjunto compacto tal que la dimensión correspondiente tiende a infinito, contradiciendo la hipótesis de que $\langle \mathbf{a}_0, \mathbf{X} \rangle$ tiene una transformada de Laplace racional. En adelante asumiremos que las dimensiones de las representaciones son siempre de orden m , las cuales podrían ser no minimales.

Sea $\mu_i(\mathbf{a}) = \frac{\mathbb{E}(\langle \mathbf{a}, \mathbf{X} \rangle^i)}{i!}$ los momentos reducidos de $\langle \mathbf{a}, \mathbf{X} \rangle$ como una función de \mathbf{a} . Entonces

$\mu_i(\mathbf{a})$ es una suma de monomios i -dimensionales en \mathbf{a} . Por el Teorema 6.1.1 tenemos que la función generadora de momentos de $\langle \mathbf{a}, \mathbf{X} \rangle$ puede ser escrita como una fracción C -continua de orden a lo más $2m$. Los determinantes de Hankel son de nuevo sumas de monomios en \mathbf{a} . Por lo tanto, colapsando la fracción continua a una función racional concluimos que la función generadora de momentos y por lo tanto su transformada de Laplace es una función racional en \mathbf{a} . El determinante de Hankel $\phi_n(\mathbf{a})$ puede eliminarse, pero a lo más en un conjunto de medida cero. La continuidad de de la transformada de Laplace multidimensional nos asegura que los coeficientes de la transformada de Laplace univariada en el conjunto nulo es obtenido como un límite de los coeficientes de la transformada de Laplace univariada fuera de este conjunto.

Se sigue que, las funciones $f_i(\mathbf{a})$ y $g_i(\mathbf{a})$ son también funciones racionales en \mathbf{a} . ■

Corolario 6.2.1: Sea $\mathbf{X} = (X_1, \dots, X_n)$ un vector que tiene una distribución matriz exponencial multivariada y sea \mathbf{A} una matriz no negativa de tamaño $m \times n$. Entonces $\mathbf{Y} = \mathbf{A}\mathbf{X}$ tiene una distribución matriz exponencial multivariada. En particular, todas las distribuciones marginales son de nuevo distribuciones matriz exponenciales.

Demostración: De acuerdo con el Teorema 6.2.1. \mathbf{Y} es una distribución matriz exponencial mutivariada si y solo si $\langle \mathbf{b}, \mathbf{Y} \rangle$ tiene una distribución matriz exponencial para todos los vectores no negativos $\mathbf{b} \neq \mathbf{0}$. Ahora $\langle \mathbf{b}, \mathbf{AX} \rangle = \langle \mathbf{bA}, \mathbf{X} \rangle$ y por lo tanto tiene una distribución matriz exponencial. ■

Teorema 6.2.2: Sea $\mathbf{X} = (X_1, \dots, X_n)$ un vector que tiene una distribución matriz exponencial multivariada. Entonces podemos escribir la función generadora de momentos para $\langle \mathbf{a}, \mathbf{X} \rangle$ como

$$\frac{\tilde{f}_1(\mathbf{a})s^{m-1} + \tilde{f}_2(\mathbf{a})s^{m-2} + \tilde{f}_{m-1}(\mathbf{a})s + 1}{\tilde{g}_0(\mathbf{a})s^m + \tilde{g}_1(\mathbf{a})s^{m-1} + \tilde{g}_{m-1}(\mathbf{a})s + 1},$$

donde los términos $\tilde{f}_i(\mathbf{a})$ y $\tilde{g}_i(\mathbf{a})$ son sumas de monomios en \mathbf{a} de orden $m-i$.

Demostración: Por Teorema 6.2.1 sabemos que la función generadora de momentos de $\langle \mathbf{a}, \mathbf{X} \rangle$ puede ser escrita como

$$\frac{f_1^*(\mathbf{a})s^{m-1} + f_2^*(\mathbf{a})s^{m-2} + f_{m-1}^*(\mathbf{a})s + f_m^*(\mathbf{a})}{g_0^*(\mathbf{a})s^m + g_1^*(\mathbf{a})s^{m-1} + g_{m-1}^*(\mathbf{a})s + g_m^*(\mathbf{a})},$$

donde $f_i^*(\mathbf{a})$ y $g_i^*(\mathbf{a})$ son sumas de monomios en \mathbf{a} de orden $m(m+1)-i$. Para ver esto, notemos que el orden de los monomios de $g_0^*(\mathbf{a})$ esta dado por la suma de los índices de la diagonal de ψ_{m+1} en (6.1.2) la cual es $m(m+1)$. Aquí $f_m^*(\mathbf{a}) = g_m^*(\mathbf{a})$ son de orden $m(m+1)-m = m^2$. La proposición del teorema es equivalente a la divisibilidad de todos los coeficientes $f_i^*(\mathbf{a})$ y $g_i^*(\mathbf{a})$ por $f_m^*(\mathbf{a})$.

Dividiendo el numerador y denominador entre $f_m^*(\mathbf{a})$ obtenemos la expresión

$$\frac{\tilde{f}_1(\mathbf{a})s^{m-1} + \tilde{f}_2(\mathbf{a})s^{m-2} + \tilde{f}_{m-1}(\mathbf{a})s + 1}{\tilde{g}_0(\mathbf{a})s^m + \tilde{g}_1(\mathbf{a})s^{m-1} + \tilde{g}_{m-1}(\mathbf{a})s + 1} = 1 + \mu_1(\mathbf{a})s + \mu_2(\mathbf{a})s^2 + \dots, \quad (6.2.1)$$

Donde $\tilde{f}_i(\mathbf{a})$ y $\tilde{g}_i(\mathbf{a})$ son ahora funciones racionales en \mathbf{a} . Esta ecuación es similar a (6.1.1) con $\alpha_{m,m} = 0$, $\alpha_{m,i} = \tilde{f}_{m-i}(\mathbf{a})$ y $\beta_{m,i} = \tilde{g}_{m-i}(\mathbf{a})$ la cual es resuelta considerando (6.1.2).

Escribimos $\tilde{g}_i(\mathbf{a}) = P_{m-i}(\mathbf{a}) + E_{m-i}(\mathbf{a})$ donde $P_{m-i}(\mathbf{a})$ es una suma de todo, si cualquier monomio de orden $m-i$ que aparece en la expresión para $\tilde{g}_i(\mathbf{a})$ mientras $E_{m-i}(\mathbf{a}) = \tilde{g}_i(\mathbf{a}) - P_{m-i}(\mathbf{a})$. Sea $\boldsymbol{\mu}_m(\mathbf{a}) = (\mu_{m+1}(\mathbf{a}), \dots, \mu_{2m}(\mathbf{a}))'$, $\phi_n = \phi_n(\mathbf{a})$ la matriz de Hankel (6.1.2) que ahora depende de \mathbf{a} , $\mathbf{P}_m(\mathbf{a}) = (P_m(\mathbf{a}), \dots, P_1(\mathbf{a}))$ y $\mathbf{E}_m(\mathbf{a}) = (E_m(\mathbf{a}), \dots, E_1(\mathbf{a}))$.

Entonces

$$-\boldsymbol{\mu}_m(\mathbf{a}) = \phi_m(\mathbf{a})\mathbf{P}_m(\mathbf{a}) + \phi_m(\mathbf{a})\mathbf{E}_m(\mathbf{a}).$$

Consideremos la j -ésima ecuación. Aquí $\boldsymbol{\mu}_{m+j}(\mathbf{a})$ es una suma de monomios de orden $m+j$ que son los términos correspondientes de $\phi_m(\mathbf{a})\mathbf{P}_m(\mathbf{a})$. Entonces $\phi_m(\mathbf{a})\mathbf{E}_m(\mathbf{a})$ es racional en \mathbf{a} y no contiene monomios de orden $m+j$, entonces concluimos al igualar coeficientes que

$$\phi_m(\mathbf{a})\mathbf{E}_m(\mathbf{a}) = \mathbf{0}$$

Entonces $\phi_m(\mathbf{a})$ es no singular y tenemos que $\mathbf{E}_m(\mathbf{a}) = \mathbf{0}$ y por lo tanto todos los términos $\tilde{g}_i(\mathbf{a})$ son sumas de monomios. De (6.2.1.) podemos ver que los términos $\tilde{f}_i(\mathbf{a})$ son sumas de monomios multiplicando ambos lados de la ecuación con el numerador e igualando coeficientes. ■

Corolario 6.2.2: El número de de momentos libres para $m, n = 2$ es a lo más

Demostración: Definimos los momentos reducidos cruzados, $\kappa_{i,j} = \frac{\mathbb{E}(X_1^i X_2^j)}{(i!j!)}$. En particular,

$\kappa_{i,0}$ y $\kappa_{0,i}$ son los momentos reducidos usuales de i -ésimo orden de X_1 y X_2 respectivamente. Entonces

$$\mu_i = \sum_{j=0}^i a_1^j a_2^{i-j} \kappa_{j,i-j}.$$

Por el Teorema (6.2.2) sabemos que $\mu_2 - \mu_1^2$ divide $\mu_3 - \mu_1\mu_2$ y $\mu_3 - \mu_1\mu_2^2$ respectivamente. Entonces existen constantes $c_{i,j}$, tal que

$$\mu_3 - \mu_1\mu_2 = (c_{1,0}a_1 + c_{0,1}a_2)(\mu_2 - \mu_1^2) \quad (6.2.2)$$

$$\mu_2 - \mu_1^2 = (c_{2,0}a_1^2 + c_{1,1}a_1a_2 + c_{0,2}a_2^2)(\mu_2 - \mu_1^2) \quad (6.2.3)$$

Igualando coeficientes en (6.2.2) tenemos

$$\kappa_{3,0} - \kappa_{1,0}\kappa_{2,0} = c_{1,0}(\kappa_{2,0} - \kappa_{1,0}^2)$$

$$\kappa_{2,1} - \kappa_{1,0}\kappa_{1,1} - \kappa_{0,1}\kappa_{2,0} = c_{0,1}(\kappa_{2,0} - \kappa_{1,0}^2) + c_{1,0}(\kappa_{1,1} - \kappa_{1,0}\kappa_{0,1})$$

$$\kappa_{0,3} - \kappa_{0,1}\kappa_{0,2} = c_{0,1}(\kappa_{0,2} - \kappa_{0,1}^2)$$

$$\kappa_{1,2} - \kappa_{0,1}\kappa_{1,1} - \kappa_{1,0}\kappa_{0,2} = c_{1,0}(\kappa_{0,2} - \kappa_{0,1}^2) + c_{0,1}(\kappa_{1,1} - \kappa_{0,1}\kappa_{1,0}).$$

Cuando $\kappa_{2,0} \neq \kappa_{1,0}^2$ y $\kappa_{0,2} \neq \kappa_{0,1}^2$ vemos que $\kappa_{1,2}$ y $\kappa_{2,1}$ están dadas únicamente en términos de otras κ 's. La ecuación (6.2.3) establece una conexión entre $\kappa_{1,2}$ y $\kappa_{2,1}$, la cual es compatible con las restricciones de (6.2.2).

Ahora, $\kappa_{2,0} = \kappa_{1,0}^2 \Rightarrow \kappa_{3,0} = \kappa_{1,0}^3$ y $\kappa_{0,2} = \kappa_{0,1}^2 \Rightarrow \kappa_{0,3} = \kappa_{0,1}^3$, lo cual completa la demostración. ■

Basado en el Teorema 6.2.1 proponemos la siguiente definición de una distribución tipo fase multivariada.

Definición: Un vector $\mathbf{X} = (X_1, \dots, X_n)$ tiene una distribución tipo fase multivariada si $\langle \mathbf{a}, \mathbf{X} \rangle$ tiene una distribución tipo fase univariada para todos los vectores no negativos $\mathbf{a} \neq \mathbf{0}$.

La siguiente definición es una extensión natural de la estructura tipo fase multivariada de las distribuciones matriz exponenciales.

Definición: Sea MME^* la subclase de las distribuciones matriz exponenciales multivariadas, donde $\langle \mathbf{a}, \mathbf{X} \rangle$ tiene representación $(\gamma, (\Delta(\mathbf{K}\mathbf{a}))^{-1} \mathbf{T}, \mathbf{t})$ donde γ , \mathbf{K} , \mathbf{T} son vectores constantes y matrices y $\mathbf{t} = -(\Delta(\mathbf{K}\mathbf{a}))^{-1} \mathbf{T}\mathbf{e}$. Decimos que $(\gamma, \mathbf{T}, \mathbf{K})$ es una representación MME^* de la distribución multivariada.

Es un problema abierto cuáles MME^* son un subconjunto propio o igual que la clase de las distribuciones matriz exponenciales multivariadas.

Teorema 6.2.3: Existen distribuciones matriz exponenciales multivariadas donde cuyo orden es estrictamente menor que las MME^*

Demostración: La prueba está basada en la inexistencia de una representación tridimensional MME^* de Krishnamoorthy y Parthasaraty exponencial multivariada para $n = 3$. La distribución es definida a través de la transformada de Laplace conjunta

$$[\mathbf{I} + \mathbf{R}\Delta(\mathbf{s})]^{-1}$$

donde \mathbf{R} es la matriz de correlación. Para encontrar un representación para $m = 3$ en $MME^*(3)$ primero parametrizamos

$$\mathbf{R} = \begin{pmatrix} 1 & \rho & \tau \\ \rho & 1 & \eta \\ \tau & \eta & 1 \end{pmatrix}$$

Entonces

$$\tilde{g}_0 = a_1 a_2 a_3 (1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2)$$

$$\tilde{g}_1 = (a_1 a_2 (1 - \rho^2) + a_1 a_3 (1 - \tau^2) + a_2 a_3 (1 - \eta^2))$$

$$\tilde{g}_2 = a_1 + a_2 + a_3$$

$$\tilde{f}_1 = 0$$

$$\tilde{f}_2 = 0.$$

Y la transformada de Laplace de $\langle \mathbf{X}, \mathbf{a} \rangle$ está dada por

$$\frac{1}{s^3 \tilde{g}_0 + s^2 \tilde{g}_1 + s \tilde{g}_2}$$

Suponemos ahora que tenemos una representación MME* $(\gamma, \mathbf{T}, \mathbf{K})$ para esta distribución en MME*. Es claro que tenemos que tener $\mathbf{K} = \mathbf{I}$. De la igualdad de la transformada de Laplace tenemos que

$$|\mathbf{T}| = \frac{-1}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$\begin{vmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{vmatrix} = \frac{1}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$\begin{vmatrix} T_{11} & T_{13} \\ T_{31} & T_{33} \end{vmatrix} = \frac{1}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$\begin{vmatrix} T_{22} & T_{33} \\ T_{32} & T_{33} \end{vmatrix} = \frac{1}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$T_{11} = \frac{1 - \eta^2}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$T_{22} = \frac{1 - \tau^2}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$T_{33} = \frac{1 - \rho^2}{1 + 2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

$$(\gamma_1 + \gamma_2) \begin{vmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{vmatrix} - \gamma_1 \begin{vmatrix} T_{12} & T_{13} \\ T_{22} & T_{23} \end{vmatrix} + \gamma_2 \begin{vmatrix} T_{11} & T_{13} \\ T_{21} & T_{23} \end{vmatrix} = 0$$

$$(\gamma_1 + \gamma_3) \begin{vmatrix} T_{11} & T_{13} \\ T_{31} & T_{33} \end{vmatrix} + \gamma_1 \begin{vmatrix} T_{12} & T_{13} \\ T_{32} & T_{33} \end{vmatrix} + \gamma_3 \begin{vmatrix} T_{11} & T_{12} \\ T_{31} & T_{32} \end{vmatrix} = 0 \quad (6.2.4)$$

$$(\gamma_2 + \gamma_3) \begin{vmatrix} T_{22} & T_{23} \\ T_{32} & T_{33} \end{vmatrix} + \gamma_2 \begin{vmatrix} T_{21} & T_{23} \\ T_{31} & T_{33} \end{vmatrix} - \gamma_3 \begin{vmatrix} T_{21} & T_{22} \\ T_{31} & T_{32} \end{vmatrix} = 0$$

$$\gamma_1 (T_{11} + T_{12} + T_{13}) = 0$$

$$\gamma_2 (T_{21} + T_{22} + T_{23}) = 0$$

$$\gamma_3 (T_{31} + T_{32} + T_{33}) = 0.$$

Debemos tener al menos una $\gamma_i \neq 0$. Debido a la simetría podemos eliminar sin pérdida de generalidad $\gamma_1 \neq 0$. Entonces tenemos que $(T_{11} + T_{12} + T_{13}) = 0$. Ahora denotemos los coeficientes de γ_j en la i -ésima ecuación de (6.2.4) por C_{ij} . Tenemos

$$C_{11} = T_{11}T_{22} - T_{12}T_{21} - T_{12}T_{23} + T_{13}T_{22}$$

$$C_{12} = T_{11}T_{22} - T_{12}T_{21} + T_{11}T_{23} - T_{13}T_{21}$$

$$C_{21} = T_{11}T_{33} - T_{13}T_{31} + T_{12}T_{33} - T_{13}T_{32}$$

$$C_{23} = T_{11}T_{33} - T_{13}T_{31} + T_{11}T_{32} - T_{12}T_{31}$$

$$C_{32} = T_{22}T_{33} - T_{23}T_{32} + T_{21}T_{33} - T_{23}T_{31}$$

$$C_{33} = T_{22}T_{33} - T_{23}T_{32} - T_{21}T_{32} + T_{22}T_{31}.$$

Sustituyendo $T_{11} = -T_{12} - T_{13}$ en las primeras cuatro ecuaciones tenemos

$$C_{11} = -T_{12}(T_{21} + T_{22} + T_{23})$$

$$C_{12} = -(T_{12} + T_{13})(T_{21} + T_{22} + T_{23})$$

$$C_{21} = -T_{13}(T_{31} + T_{23} + T_{33})$$

$$C_{23} = -(T_{12} + T_{13})(T_{31} + T_{32} + T_{33}).$$

Como $\gamma_1 C_{12} = \gamma_3 C_{23} = 0$ tenemos que

$$\gamma_1 C_{11} = \gamma_1 T_{12} (T_{21} + T_{22} + T_{23}) = 0$$

$$\gamma_1 C_{21} = \gamma_1 T_{13} (T_{31} + T_{23} + T_{33}) = 0.$$

No podemos tener T_{12} y T_{13} iguales a cero y tampoco podemos tener $T_{21} + T_{22} + T_{23} = 0$ y $T_{31} + T_{23} + T_{33} = 0$. Debido a la simetría podemos suponer sin pérdida de generalidad que $T_{13} = 0$ y $T_{21} + T_{22} + T_{23} = 0$ mientras que $T_{12} \neq 0$ y $T_{31} + T_{23} + T_{33} \neq 0$; por lo tanto

$$C_{32} = -T_{23}(T_{31} + T_{23} + T_{33})$$

$$C_{33} = -(T_{21} + T_{23})(T_{31} + T_{23} + T_{33}).$$

Entonces $T_{21} + T_{22} + T_{23} = 0$ tenemos que $\gamma_3 = 0$ y no podemos tener T_{13} y T_{23} igual a cero, entonces concluimos que $\gamma_2 = 0$. Finalmente tenemos que

$$T_{11}T_{33} = \frac{T_{11}}{1-\eta^2} = \frac{T_{33}}{1-\rho^2} = \frac{1}{1+2\rho\tau\eta - \rho^2 - \tau^2 - \eta^2}$$

La cual sólo se cumple cuando $\tau = \rho\eta$. Podemos encontrar una representación en las distribuciones tipo fase multivariadas con

$$\mathbf{T} = \begin{pmatrix} -\frac{1}{1-\rho^2} & \frac{1}{1-\rho^2} & 0 \\ \frac{\rho^2}{1-\rho^2} & -\frac{1-\eta^2\rho^2}{(1-\rho^2)(1-\eta^2)} & \frac{1}{1-\eta^2} \\ 0 & \frac{\eta^2}{1-\eta^2} & -\frac{1}{1-\eta^2} \end{pmatrix}$$

Es posible obtener una representación en MME* siempre que una de las tres ecuaciones $\tau = \rho\eta$, $\rho = \tau\eta$, o $\eta = \rho\tau$ es cierta. ■

6.3 Distribuciones Matriz Exponenciales Multivariadas y Teoría de Renovación

Mostraremos las aplicaciones de las distribuciones matriz exponenciales multivariadas como la teoría de renovación.

Lema 6.3.1: Sea A_t y R_t la edad y el tiempo que resta de vida respectivamente en un proceso de renovación estacionario con función de densidad del tiempo entre arribos $f(x)$. La distribución conjunta de A_t y R_t esta dada por $\frac{f(x+y)}{\int_0^\infty xf(x)dx}$.

Demostración: La distribución conjunta de (A_t, R_t) en un proceso de renovación está dada por

$$\mathbb{P}(R_t \geq x, A_t < y) = \int_{t-y}^t (1-F(t+x-u))dU(u),$$

Obtenemos el resultado, insertando $\frac{1}{\mathbb{E}(X)}$ como la densidad de renovación estacionaria y diferenciando. ■

Teorema 6.3.1: Un proceso de renovación estacionario con tiempos entre arribos distribuidos matriz exponencial con representación (\mathbf{a}, \mathbf{C}) , la distribución conjunta de edad y vida residual es una MME* con representación

$$\left(\left(\frac{\mathbf{a}(-\mathbf{C})^{-1}}{\mu_1}, \mathbf{0} \right), \begin{pmatrix} \mathbf{C} & -\mathbf{C} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}, \begin{pmatrix} \mathbf{e} & \mathbf{0} \\ \mathbf{0} & \mathbf{e} \end{pmatrix} \right), \quad (6.3.1)$$

Donde m es la dimensión de \mathbf{C} .

Demostración: Denotemos la densidad de la distribución matriz exponencial con representación (\mathbf{a}, \mathbf{C}) por $f(x)$. La media correspondiente es denotada por μ_1 . Usando el Teorema (6.2.1) observamos que para $a_1 > 0$ y $a_2 > 0$, la variable aleatoria $Z = \langle \mathbf{a}, \mathbf{X} \rangle$, tiene densidad $g(z)$ dada por

$$g(z) = \frac{1}{\mu_1} \int_0^{\frac{x}{a_1}} \mathbf{a} \exp \left(\mathbf{C} \left(x_1 + \frac{z - a_1 x_1}{a_2} \right) \right) c dx_1 \frac{1}{a_2}.$$

Para $a_1 \neq a_2$ tenemos

$$g(z) = \frac{1}{\mu_1} \mathbf{a} \mathbf{C}^{-1} \frac{1}{1 - \frac{a_1}{a_2}} \left[\exp \left(\mathbf{C} \frac{z}{a_1} \right) \exp \left(\mathbf{C} \frac{z}{a_2} \right) - \mathbf{I} \right] \exp \left(\mathbf{C} \frac{z}{a_2} \right) \frac{c}{a_2}$$

$$g(z) = \frac{1}{\mu_1} \frac{1}{a_2 - a_1} \left(F \left(\frac{z}{a_1} \right) - F \left(\frac{z}{a_2} \right) \right)$$

la cual es también la densidad de una distribución matriz exponencial con la representación dada por (6.3.1). Esto puede ser visto evaluando directamente la cola. Para $a_1 = a_2$ tenemos

$$g(z) = \frac{z f \left(\frac{z}{a} \right)}{a \mu_1}. \quad \text{Otra vez por verificación directa, también obtenemos la densidad evaluando}$$

(6.3.1) para $a_1 = a_2$ ■

En este capítulo se analizó la clase de las distribuciones matriz exponenciales multivariadas definidas como distribuciones con transformada de Laplace racional multidimensional, se mostró el teorema de caracterización principal, así como resultados acerca del orden de las distribuciones matriz exponenciales multivariadas.

Conclusiones

El propósito de este trabajo fue mostrar los resultados acerca de las distribuciones matriz exponenciales y aprovechar la correspondencia uno a uno que existe entre estas y las distribuciones con transformada de Laplace racional, para poder describir los modelos estocásticos de flujos como distribuciones matriz exponenciales y determinar la probabilidad de ruina en un modelo de riesgo con distribuciones matriz exponenciales.

Por otro lado mostrar que las distribuciones matriz exponenciales ofrecen soluciones algorítmicamente tratables.

Se definieron algunos conceptos preliminares que fueron utilizados a lo largo de este trabajo debido a que basamos gran parte de la teoría en éstos. Algunos de los conceptos más interesantes son los procesos de riesgo y el proceso de superávit de reclamaciones que son modelos para la evolución de la reserva de una compañía aseguradora en el tiempo; también se describieron los principales métodos que se utilizan en la teoría del riesgo aunque a lo largo del presente trabajo nos enfocamos en los métodos analíticos matriciales. Se definió formalmente la transformada de Laplace y se describieron sus propiedades principales. Además fueron descritas las ecuaciones diferenciales de Kolmogorov y por último se describieron algunos resultados de álgebra matricial que fueron utilizados en los capítulos posteriores.

Se describieron los procesos de Markov que son modelos que describen situaciones donde los estados cambian en puntos del tiempo arbitrarios. Estos procesos son gobernados por una matriz de intensidad la cual es única para la distribución del proceso de Markov minimal. También se mostraron los resultados correspondientes a las ecuaciones diferenciales de Kolmogorov para los procesos de Markov, la clasificación de los estados y las medidas estacionarias de estos procesos.

Se mostraron las distribuciones tipo fase como la distribución del tiempo de paro hasta la absorción $\eta = \inf\{t \geq 0 \mid X_t = 0\}$ para un proceso de Markov finito de saltos $\{X_t\}$ con espacio de estados $E' = \{1, \dots, m+1\}$ en el caso univariado y multivariado, se mostraron sus propiedades, representaciones y caracterizaciones; así como también las propiedades de clase de las distribuciones tipo fase. Además se analizó un proceso de riesgo donde las reclamaciones se distribuyen tipo fase. Finalmente se mostraron las distribuciones tipo fase multivariadas y su caracterización.

Se analizaron las distribuciones matriz exponenciales y dada la correspondencia uno a uno entre la transformada de Laplace racional y las representaciones de las distribuciones matriz exponenciales se obtuvieron resultados importantes en términos de los vectores \mathbf{a} y \mathbf{b} . También se mostraron los procesos de renovación donde la distribución entre arribos \mathbf{B} es matriz exponencial y el método que permite escribir $\mathbf{T} + \mathbf{a}\mathbf{s}$ de forma diagonal y directamente obtener la función matriz exponencial necesitada para la densidad de renovación.

Se mostró la interpretación física que Bladt y Neuts (2003) ofrecen acerca de las distribuciones matriz exponenciales definiendo el flujo válido para poder definir la relación entre Y , que es una variable aleatoria con distribución matriz exponencial (\mathbf{a}, \mathbf{S}) , $T(U)$ donde U se distribuye como una variable aleatoria uniforme $[0,1]$. Además se mostraron resultados importantes sobre los procesos de renovación con tiempos de arribo matriz exponencial modelados como secuencias de flujos terminales y el modelo de riesgo que Asmussen y Bladt (199) modelaron bajo la interpretación física de flujos. Por último se mostró la interpretación física para los procesos de arribo racionales con una representación de órbitas.

Se analizó la clase de las distribuciones matriz exponenciales multivariadas definidas como distribuciones con transformada de Laplace racional multidimensional. Esta clase de distribuciones generaliza clases de distribuciones exponenciales y gamma multivariadas.

Se mostró el teorema de caracterización principal, el cual demuestra que las distribuciones matriz exponenciales multivariadas que son proyecciones en cualquier dirección son distribuciones matriz exponenciales multivariadas. Basado en la teoría de fracciones continuas se probó un resultado importante correspondiente al orden de las distribuciones matriz exponenciales y la desaparición del determinante de Hankel de los momentos reducidos.

A lo largo de este trabajo hemos podido observar las ventajas que presentan las distribuciones matriz exponenciales en los modelos estocásticos que las utilizan al tener mayor flexibilidad y generalidad. El resultado más interesante en este trabajo es sin duda alguna la interpretación física de las distribuciones matriz exponenciales y la interpretación matemática de los flujos válidos para simplificar el análisis de las distribuciones matriz exponenciales y obtener importantes resultados en teoría de renovación y aplicar estos a distintas áreas particularmente en la teoría del riesgo y así poder calcular probabilidades de ruina.

Existe material bibliográfico que profundiza aún más, acerca de las distribuciones matriz exponenciales donde se describe la estructura del espacio de probabilidad Ω_p de las distribuciones matriz exponenciales de dimensión p , algoritmos para identificar distribuciones matriz exponenciales, así como también el ajuste de datos a distribuciones matriz exponenciales por medio de la estimación de máxima verosimilitud (Ver Fackrell 2003).

Actualmente existe poco material bibliográfico acerca de las distribuciones matriz exponenciales multivariadas y sus posibles aplicaciones; por este motivo es importante dirigir la investigación a las aplicaciones de las distribuciones matriz exponenciales multivariadas en distintas áreas.

Bibliografía

- Asmussen, S. *Applied Probability and Queues*. Springer-Verlag New York, U.S.A., Second Edition, 2003.
- Asmussen, S. *Ruin Probabilities*, Volume 2 of *Advanced Series on Statistical Science and Applied Probability*. World Scientific Publishing Company, Singapore, 2000.
- Asmussen, S. and Bladt, M. *Renewal Theory and Queueing Algorithms for Matrix-Exponential Distributions*. In Chakravarthy, S. R. and Alfa, A.S., editors. *Matrix-Analytic Methods in Stochastic Models*, Volume 183 of *Lecture Notes in Pure and Applied Mathematics*, pages 313-341. Marcel Dekker, New York, U.S.A., 1997.
- Asmussen, S. and Bladt, M. Point processes with finite-dimensional conditional probabilities. *Stochastic Processes and their Applications*; 82, pages 127-142, 1999.
- Assaf, D., Langberg, N.A., Savits, T.H, Shaked M. *Multivariate Phase-Type Distributions*. *Operations Research*, 32(3), pages 688-702, 1984.
- Bladt, M. *Teoría del Riesgo I. Apuntes de Clase*. México, 2004.
- Bladt, M. and Neuts, M.F. *Matrix-Exponential Distributions: Calculus and Interpretations via Flows*. *Communications in Statistics-Stochastic Models* Vol. 19, No. 1; pages 113-124. Marcel Dekker, New York, U.S.A., 2003.
- Bladt, M, and Nielsen B.F. *Multivariate Matrix-Exponential Distributions*. Technical University of Denmark, Department of Informatics and Mathematical Modeling. Denmark, 2008.
- Erlang, A: K. *Solution of some problems in the teory of probabilities of significance in automatic telephone exchanges*. *The Post Office Electrical Engineer's Journal*, 10, pages 189-197. 1917.
- Fackrell, M.W. *Characterization of Matrix-Exponential Distributions*. *PhD Thesis*. School of Applied Mathematics. The University of Adelaide, Adelaide, Australia, 2003.
- Grimmett, G.R. and Stirzaker, D.R. *Probability and Random Processes*. Oxford University Press, Oxford, England, Third Edition, 2001.
- Kulkarni, V.G. A new class of multivariate phase type distributions. *Operations Research*, 37(1), pages 151-158, 1989.
- Latouche, G. and Ramaswami, V. *Introduction to Matrix Analytic Methods in Stochastic Models*. ASA-SIAM, Philadelphia, U.S.A, 1999.
- Neuts, M. F. *Matrix-Geometric Solutions in Stochastic Models: an Algorithmic Approach*; Dover Publications, Inc., New York, U.S.A., 1981.
- Neuts, M.F. Probability distributions of phase type. *Liber Amicorum Prof. Emeritus. H. Florin*, pages 173-206, Department of Mathematics, University of Luvain, Luvain, Belgium, 1975.
- Perron, O. *Die lehre von den Kettenbrüchen*. B.G. Teubner Verlagsgesellschaft, Stuttgart, Deutschland, 1957.
- Rolski, T., Schmidli, H., Schmidt, V., Teugels, J. *Stochastic Processes for Insurance and Finance*. John Wiley & Sons, Chichester, England, 1999.

- Van de Liefvoort, A. *The moment problem for continuous distributions*. Technical Report WP-CM-1990-02, University of Missouri, Kansas City, U.S.A., 1990.