



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**UNA METODOLOGÍA PARA COMPARAR
MODELOS DE RIESGO DE CRÉDITO**

T E S I S

QUE PARA OBTENER EL TÍTULO DE :

A C T U A R I O

P R E S E N T A :

FERNANDO BALTAZAR LARIOS



DIRECTOR DE TESIS : DR. PABLO PADILLA LONGORIA



m340342

**FACULTAD DE CIENCIAS
SECCION ESCOLAR**



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.



SECRETARÍA DE EDUCACIÓN PÚBLICA
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: BALTAZAR LARIOS FERNANDO

FECHA: 25/01/05

FIRMA: [Firma]

ACT. MAURICIO AGUILAR GONZÁLEZ
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

"Una metodología para comparar modelos de riesgo de crédito"

realizado por Fernando Baltazar Larios

con número de cuenta 400000411 , quien cubrió los créditos de la carrera de:
Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis

Propietario

Dr. Pablo Padilla Longoria

Pablo Padilla L.

Propietario

Dr. Mogens Bladt Petersen

Mogens Bladt

Propietario

Dr. Ramsés Humberto Mena Cáravez

[Firma]

Suplente

Dra. Ana Meda Guardiola

Ana Meda

Suplente

Dr. Luis Antonio Rincón Solís

[Firma]

Consejo Departamental de Matemáticas

[Firma]
Act. Jaime Vázquez Alamilla
Coordinador de la carrera
de Actuaría

2005/01/25

Agradecimientos

Agradezco a Dios porque sin él nunca lo hubiera logrado

A mis padres y a mis hermanas por todo el amor que me han dado.

A Karina por la paciencia y amor brindados.

Al Dr. Pablo Padilla Longoria por todo el apoyo que me brindo para el desarrollo de mi tesis.

A mis sinodales Dra. Ana Meda Guardiola, Dr. Ramses Mena Chávez, Dr. Mogens Bladt y Dr. Luis Antonio Rincón Solís por su comprensión y ayuda para realizar este trabajo.

Dedicatorias

A ti, papá, por el ejemplo que me has dado, por todas las cosas que me has enseñado y sobre todo el amor que das.

A ti, mamá, por ser la mujer que eres, por estar conmigo, por brindarme el apoyo que siempre he necesitado y por tanto que me has amado.

A ti, mi hermana, Marcela, por haber sido el mejor ejemplo de un buen hijo que pude tener a mi lado, por todas las alegrías que hemos vivido y el amor que siempre me has dado.

A ti, mi hermanita, Mariana, por la infancia que sólo tú me pudiste haber dado, por toda la felicidad que hemos compartido y el amor que a tu lado nunca me ha faltado.

A ti, Karina, por todas las cosas preciosas que me has enseñado, por todo el amor que me has dado, por la felicidad que me has otorgado y porque se que siempre estarás a mi lado.

Quiero que sepan que aunque no podíamos estar juntos, siempre estuvieron en mi mente, pero sobre todo en mi corazón.

A ustedes, mis amigos con quienes todo he compartido, Queso, que siempre

has sido el hermano que he necesitado, Polivoz, Malilla, Chamel, Güero, Chiquito, Chango y Arnold, por todos los momentos que hemos pasado y la amistad que me han dado.

Al Kike, a su familia y el Rafa por el cariño que desde el primer momento me brindaron.

A mis amigos del chilango que sin su cariño jamás lo hubiese logrado, Jos, Filos, Rice, Galileo, Benjitas, Charly, Chay y Helen.

Al Nacho, por cuidar de mi hermana y mi familia.

Y muy en especial quiero dedicar este trabajo a mis sobrinos, los Toros, quiero que sepan que fueron la inspiración y que los amo con todo mi corazón.

Contenido

1. Riesgo de Crédito	11
1.1. Riesgo	11
1.1.1. Tipos de Riesgo	11
1.1.2. Tipos de Pérdidas	14
1.2. Crédito	14
1.3. Riesgo de Crédito	15
1.3.1. Importancia del Riesgo de Crédito	16
1.3.2. Límites	19
1.4. Modelos de Riesgo de Crédito	19
1.4.1. La Calificación de Riesgo de Crédito	20
1.4.2. Técnicas Empleadas en la Construcción de Modelos	21
1.4.3. Clasificación de los Modelos	22
2. Métodos de Entropía	37
2.1. Introducción	37
2.2. Máxima Entropía	38
2.2.1. El Problema	38
2.2.2. Medida de Entropía	39
2.2.3. Entropía Normalizada	47
2.3. Mínima Entropía Relativa	49
2.4. Entropía en Variables Aleatorias	52
2.4.1. Variables Aleatorias Discretas	52
2.4.2. Variables Aleatorias Continuas	54
3. Matrices de Migración de Crédito	57
3.1. Calificaciones Crediticias	57
3.1.1. Compañías Calificadoras	58
3.1.2. Importancia de las Calificaciones de Riesgo de Crédito	61

3.2. Matrices de Transición	62
3.2.1. Cadenas de Markov y Matrices de Transición	63
3.2.2. La Matriz de Transición de Calificaciones Crediticias	67
3.3. Métodos de Estimación de Matrices de Transición	69
3.3.1. Modelo Regular de Migración de Crédito	69
3.3.2. Método <i>Cohort</i>	81
3.3.3. Método de Duración o Tasa de Riesgo	81
3.3.4. Máxima Entropía	82
3.3.5. Matrices Potencia	83
4. Comparación de Matrices de Transición de Riesgo de Crédito	85
4.1. Introducción	85
4.2. Aplicaciones de la Entropía	86
4.2.1. Entropía en la Física	86
4.2.2. La Entropía en la Teoría de la Comunicación	86
4.2.3. La Entropía en Economía	89
4.3. Generación de Matrices de Transición Crediticias	91
4.3.1. La Motivación	91
4.4. Comparación de Matrices de Transición de Riesgo Crediticio con Entropía Relativa	95
4.4.1. Minimizando la Entropía Relativa	95
4.4.2. Aplicación	96
5. Conclusiones	99
A. Simulación Uniforme	101
B. Simulación Normal	105
C. Método OM	109
D. Método OMG	111
E. Matrices Potencia	113

Introducción

Debido a la importancia que ha tomado el riesgo de crédito en los últimos años por diversos factores, además de las exigencias en cuanto a regulación de los boletines emitidos por el comité de Basilea para la estimación del riesgo de crédito, resulta de gran importancia una óptima estimación de éste.

Motivados en lograr la mejor estimación posible del riesgo de crédito basados en los diferentes modelos propuestos para su estimación, en este trabajo se propone un método de comparación de modelos de riesgo de crédito. Es decir, se desarrolla una metodología, basado en la producción de mínima entropía, para comparar los diferentes modelos de riesgo de crédito que permite escoger el que realice la mejor estimación de este.

Con esta finalidad a lo largo de este trabajo se realizó lo siguiente:

En el primer capítulo se exponen los conceptos fundamentales relacionados con el riesgo de crédito. Se discute su importancia y sobre todo las consecuencias que se derivan de una mala estimación del mismo. Se estudian los diferentes modelos de riesgo de crédito que han sido más importantes y se expone un panorama actual general del mundo del riesgo crediticio.

En el segundo capítulo se presenta el concepto de medida de incertidumbre, se desarrolla el concepto de máxima entropía y de mínima entropía relativa. De igual forma se presentan los fundamentos para aplicar dichos conceptos a la estimación del riesgo de crédito.

En el tercer capítulo se presenta el papel que desempeña una calificación de riesgo de crédito en el mundo financiero. También se desarrolla el concepto de matriz de transición de riesgo de crédito y algunos métodos que permiten estimarlas.

En el cuarto capítulo se realiza la estimación de matrices de transición de riesgo crediticio y con base en éstas se compara los correspondientes modelos de riesgo de crédito, tomando como criterio de selección los conceptos de entropía desarrollados en el segundo capítulo. También se presentan algunas otras aplicaciones relevantes de dichos conceptos y la importancia que tiene formular un criterio de esta índole en el mundo financiero. Concluyendo con la selección de un método de estimación de calificación crediticia.

Capítulo 1

Riesgo de Crédito

En este capítulo se definen el riesgo, crédito y el riesgo de crédito. Se describe la importancia del riesgo de crédito y las principales herramientas usadas para su evaluación. Además, se presenta una descripción de la clasificación de los principales modelos para estimar el riesgo de crédito.

1.1. Riesgo

El riesgo se entiende como la posible **pérdida financiera** por los distintos factores aleatorios presentes en el **mercado**.

La rentabilidad debe ser proporcional al riesgo: **a mayor rentabilidad, mayor riesgo**, y viceversa. Si no se quiere correr **ningún riesgo**, se debe invertir en instrumentos financiero libres de riesgo (los emitidos por el gobierno federal) y se obtendrá la rentabilidad libre de riesgo. Si se invierte en un activo con riesgo, se espera obtener la rentabilidad libre de riesgo más una prima de rentabilidad o prima de riesgo.

1.1.1. Tipos de Riesgo

La rentabilidad de un título (empresa, entidad o portafolio) está afectada por dos tipos de riesgo:

1. Un riesgo propio o específico que depende de las características propias del título o empresa emisora (naturaleza de sus actividades productivas, competencia de la gerencia, etc.). A este riesgo también se le conoce como no sistemático o diversificable.
2. El otro, sistemático o de mercado, que no depende de las características intrínsecas del título, sino de la estructura del mercado (tasas de interés, tipo de cambio, etc). A este riesgo no se le puede eliminar mediante la diversificación dada la correlación existente entre la rentabilidad del título en cuestión con las rentabilidades de otros títulos por medio del índice bursátil que resume la evolución del mercado.

Si un inversionista compra títulos en el mercado de valores con el fin de reducir el riesgo, entonces, tiene sentido la diversificación si las rentabilidades de los diferentes títulos adquiridos no están correlacionadas, o tienen distinto grado de correlación con el índice de mercado.

El modelo más conocido para estimar la rentabilidad y el riesgo de los valores mobiliarios es el modelo de mercado de *Sharpe*¹. Éste sirve de base al modelo diagonal, en el cual se parte de la dependencia estadística de tipo lineal existente entre la rentabilidad de los títulos y la del mercado en general.

Uno de los criterios para la clasificación de los activos financieros es el basado en el coeficiente *Beta de Sharpe*[9] o coeficiente de volatilidad. La *Beta* es un indicador de riesgo, marca en qué medida un título sigue las fluctuaciones del mercado.

La medida del riesgo de mercado viene dada por la *Beta*. Ésta vincula los retornos del mercado con los de una inversión en particular.

Según el coeficiente *Beta de Sharpe* los activos financieros se clasifican en:

1. Activos poco volátiles o defensivos ($beta < 1$).
2. Activos muy volátiles o agresivos ($beta > 1$).
3. Activos de volatilidad normal o neutros ($beta = 1$).

¹En este modelo *Sharpe* propone relacionar la evolución de la rentabilidad de cada activo financiero con un determinado índice, normalmente macroeconómico

El coeficiente *Beta* indica la respuesta de la rentabilidad de la acción al riesgo sistemático, y mide la sensibilidad de la rentabilidad de un título al factor de riesgo específico (rentabilidad del mercado).

Si, por ejemplo, las acciones de una empresa se relacionan positivamente con el riesgo de la inflación, tales acciones tienen una *Beta* de inflación positiva. Si se relacionan negativamente con la inflación, su *Beta* de inflación es negativa, y si no se correlacionan con la inflación, su *Beta* de inflación es cero.

Ejemplo:

Supongamos que se han identificado dos riesgos sistemáticos que necesitamos tener en cuenta y que son suficientes para describir los riesgos sistemáticos que influyen en la rentabilidad de las acciones. Éstos son: inflación y tasas de interés. Todas las acciones tendrán una *Beta* asociada con uno de estos riesgos sistemáticos: Una *Beta* de inflación y una *Beta* de la tasa de interés.

Supongamos que a principios de año se pronostica una inflación de 5% y que la tasa de interés permanecerá constante. Las acciones que se contemplan tienen las siguientes *Betas*:

$$I = 2 \text{ y } r = -1.8$$

Ahora supongamos que durante el año han ocurrido los siguientes hechos:

1. La inflación aumenta el 7%.
2. La tasa de interés baja el 2%.

Entonces tenemos que:

$$F_I = \text{Inflación Real} - \text{Inflación Esperada} = 7\% - 5\% = 2\%.$$

$$\text{y}$$

$$F_r = \text{Cambio real} - \text{Cambio esperado} = -2\% - 0\% = -2\%.$$

También sabemos que:

$$m = \text{Parte del riesgo sistemático de la rentabilidad} = IF_I + rF_r$$

$$= (2 \times 2\%) + ((-1.8) \times (-2\%)) = 7.6\%$$

Si, por ejemplo, la rentabilidad esperada de las acciones para el año es del 4%, la rentabilidad total de los dos componentes será:

$$R = 4\% + 7.6\% = 11.6\%$$

En el ejemplo anterior es utilizado *Modelo de Factor* [15], y las fuentes sistémicas de riesgo, que denominamos *F*, reciben el nombre de Factores.

1.1.2. Tipos de Pérdidas

Las pérdidas pueden ser clasificadas en dos tipos:

1. Las que son estimadas por medio de modelos matemáticos, llamadas pérdidas esperadas.
2. Las que no es posible estimar debido a que las determinan variables exógenas ², conocidas como pérdidas no esperadas.

1.2. Crédito

Es un contrato en el cual existe una transferencia de dinero que debe ser devuelto después de un tiempo determinado. En este contrato hay varios elementos. El acreedor que es quien transfiere el dinero, el deudor que es quien lo recibe y los términos del crédito.

Los créditos se pueden clasificar en dos tipos por su tiempo de vencimiento (convención financiera) :

1. A corto plazo, que son los créditos que vencen en un período menor de un año.
2. A largo plazo, son créditos cuyo vencimiento se presenta después de un año.

²Es una variable cuyo valor se determina fuera del modelo en el cual se está trabajando.

1.3. Riesgo de Crédito

El riesgo de crédito es el riesgo asociado a la posibilidad de que la contraparte, en un contrato de crédito, no cumpla con las obligaciones derivadas del mismo y por tanto cause a sus acreedores una pérdida financiera. Cabe destacar que bajo esta definición es irrelevante si la contraparte no puede o no quiere cumplir con su obligación.

El riesgo de crédito puede ser analizado en tres dimensiones básicas:

1. El Riesgo de Incumplimiento.

Es la probabilidad de que se presente un incumplimiento en el pago de un crédito. Dicho incumplimiento podría ser puramente económico, es decir, que no está asociado a ningún evento específico. Esto ocurre cuando el valor económico de los activos se reduce por debajo del saldo remanente del adeudo.

Es importante estimar la probabilidad de incumplimiento, pues aunque el simple hecho de un incumplimiento parcial no genera pérdidas inmediatas, sí se incrementa la probabilidad de un incumplimiento total.

Este riesgo es medido por el cálculo de la probabilidad de que ocurra un incumplimiento en un periodo dado de tiempo.

El riesgo de incumplimiento depende de dos aspectos:

- a) Aspectos internos al acreditado, como la situación crediticia del mismo.
- b) Aspectos externos, tales como la situación económica del país o el comportamiento de los mercados financieros nacionales e internacionales.

Además, es importante estimar la tasa de deterioro (*roll rate*), es decir, cuántos de los acreditados que incumplen una, dos o tres veces llegarán al incumplimiento total.

2. Riesgo de Exposición.

Es el generado por la incertidumbre respecto a los montos futuros en riesgo. Cuando los créditos pueden pagarse total o parcialmente de manera anticipada, se presenta el riesgo de exposición, ya que al no cumplirse con exactitud el plazo de liquidación se dificulta la estimación de los montos en riesgo.

Por ejemplo, los productos derivados tradicionales generan riesgo de exposición, pero en este caso, el factor de incertidumbre no se relaciona con el comportamiento del acreditado, por lo que el valor de liquidación del derivado depende de las constantes fluctuaciones del mercado.

3. El Riesgo de Recuperación.

Este riesgo está presente en el evento de un incumplimiento, ya que la recuperación es impredecible por depender del tipo de incumplimiento y de factores relacionados con las garantías que se hayan recibido, el tipo de garantía y su situación al momento de incumplimiento.

Con los avales se modifica el riesgo de crédito ya que en caso de incumplimiento se traslada la incertidumbre del acreditado al aval. Sin embargo, ésta no es una simple transferencia de riesgo, pues si el acreditado y su aval incumplieran al mismo tiempo, la probabilidad correspondiente sería una probabilidad conjunta por incumplimiento.

Además, para estimar la recuperación, es necesario que se consideren los aspectos legales que ésta pudiera implicar (proceso de reconocimiento de adeudo, proceso de elaboración de documentos necesarios para tomar acción legal, probabilidad de que la acción legal no sea exitosa, etc.) entonces, la acción de recuperación involucra el riesgo legal.

1.3.1. Importancia del Riesgo de Crédito

El riesgo de crédito es el de mayor importancia en cuanto a las pérdidas potenciales que su inadecuado manejo puede implicar para una institución de crédito.

En años recientes han surgido nuevas técnicas en la forma de medir y administrar el riesgo de crédito, principalmente siguiendo las siguientes líneas:

1. Medidas de riesgo de crédito para cartera de créditos.
2. Medidas de riesgo de crédito para los instrumentos de mercado.
3. Determinación del VaR³ para riesgo de crédito.
4. Surgimiento de técnicas para la administración de la cartera de crédito.

Los principales fenómenos que han originado un sorpresivo surgimiento del interés sobre el tema de riesgo de crédito a nivel mundial son:

1. Crecimiento estructural de las deudas.

Considerando el argumento estructural y significativo de las quiebras en el mundo y dadas las especificaciones de las estadísticas comparativas de quiebras actuales con las de crisis previas que señalan un crecimiento significativo, es importante poner mayor atención en la estimación del riesgo de crédito. Además el aumento en la competencia global, hacen hoy en día más importante un estricto y adecuado análisis de crédito que en el pasado.

2. Desintermediación.

Debido a la expansión de los mercados de capitales, los créditos se han vuelto más accesibles a empresas pequeñas y medianas. Así mismo, está disminuyendo el número de empresas rezagadas en la obtención de fondos de bancos y de otras instituciones financieras, a pesar de que cuentan con calificaciones crediticias más débiles.

3. Márgenes más competitivos.

No obstante la caída en la calidad promedio de los créditos (la capacidad promedio de los acreditado para cumplir con sus compromisos en un contrato de crédito), los márgenes o *spreads*⁴ se han vuelto muy estrechos, es decir, que el riesgo inherente a la actividad de prestar ha aumentado.

³Valor en Riesgo

⁴Es la diferencia porcentual entre las tasas de los bonos libres de riesgo y el corporativo.

4. Disminución del valor los activo físicos (bienes inmuebles) y la volatilidad de su precio.

Debido a que cuanto más débil e incierto es el valor de estos, más débil se vuelve la actividad de préstamo, se han incrementado las preocupaciones sobre el valor de las propiedades y otros activos físicos.

5. El crecimiento de derivados fuera de balance ha ocasionado el crecimiento del riesgo de crédito y se ha extendido la necesidad del análisis de crédito mas allá de lo tradicional.
6. Los avances en la tecnología le han dado a los bancos y a las instituciones financieras tradicionales las herramientas para trabajar con técnicas de elaboración de modelos altamente tecnificados.
7. Posiblemente el mayor impulso para que, tanto los bancos, como las instituciones financieras se hayan interesado en el desarrollo de nuevos modelos de riesgo de crédito fue su inconformidad con las medidas impuestas por el *BIS*⁵.

El riesgo de crédito debe ser medido en dos niveles.

1. Análisis crediticio para el otorgamiento del crédito, el cual incluye los siguientes aspectos a analizar:
 - a) Las razones por las que el cliente requiere el credito.
 - b) Los estados financieros y su tendencia.
 - c) Las razones financieras.
 - d) Los fines con los que será utilizado el crédito y su impacto en los flujos de caja.
 - e) La sensibilidad para estimar la fortaleza de las proyecciones, es decir. conocer el comportamiento pasado crediticio del cliente para estimar su comportamiento futuro.
 - f) El sector económico al que pertenece el cliente.

⁵Bank International Settlements (Basilea).

- g) Evaluación de la administración de la empresa y su estrategia operativa.
 - h) Establecer restricciones operativas, financieras, garantías y avales.
2. Análisis del riesgo de crédito de cartera, tal y como se define en la institución financiera reguladora correspondiente (CNBV ⁶ circular 1473).

1.3.2. Límites

Es importante y necesario que los bancos y las instituciones financieras establezcan límites de crédito para acotar las pérdidas en caso de incumplimiento. Debe existir un proceso de autorización para establecer el monto máximo en riesgo que se está dispuesto a asumir ante un determinado cliente antes del otorgamiento de cualquier crédito.

Los objetivos de establecer límites son:

1. Evitar que la pérdida en un sólo crédito ponga en riesgo a la institución financiera.
2. Diversificación de compromisos de otorgamiento de crédito en varias dimensiones (por cliente, sector económico, zona geográfica, etc).
3. No otorgar crédito a cualquiera por un monto que exceda su capacidad de endeudamiento.

1.4. Modelos de Riesgo de Crédito

Los modelos de riesgo de crédito calculan la probabilidad de incumplimiento en el pago de los créditos otorgados y cuantifican el riesgo al enfrentar este incumplimiento.

Existen varios factores que afectan la probabilidad de incumplimiento, tales como, las tasas de interés, la variación en el precio de insumos, la competencia del mercado, etc.; los modelos buscan anticipar estos eventos y con ello determinar el riesgo de crédito con mayor precisión.

⁶ Comisión Nacional Bancaria y de Valores

Para diseñar un modelo se necesita partir del establecimiento de las relaciones entre las diversas variables que afectan al riesgo y partiendo de ellas construir el modelo.

1.4.1. La Calificación de Riesgo de Crédito

La calificación de riesgo tuvo su origen a finales del siglo antepasado en Estados Unidos con la consolidación de un sistema de información crediticia, mismo que era utilizado por inversionistas e instituciones financieras de ese país. A principios de el siglo pasado, con el desarrollo de la industria ferroviaria y con la importancia que obtuvo la emisión de bonos para su financiamiento, se crearon nuevas empresas calificadoras para estudiar la calidad de estos valores. Varias de las firmas más importantes de los EE.UU. dieron sus primeros pasos en el campo de la evaluación de riesgo, por ejemplo *Poor's Publishing Co* publicó su primera calificación en 1916 y *Standard Statistics Bureau* inicia actividades en 1922. Sólo hasta 1972 se funda la primera calificadora fuera de los EE.UU., específicamente en Canadá. En Europa la primera de estas firmas surgió en España en 1985, con el nombre *Renta 4 S.A.*

En América Latina la primera calificadora de riesgo se autorizó en Chile en 1988 y el segundo país de la región en donde se constituyó fue en México, en enero de 1990.

La calificación de riesgo es la opinión de una entidad independiente especializada en estudios de riesgo sobre la calidad crediticia de una emisión de valores. La evaluación se realiza sobre la capacidad de la entidad emisora de cumplir puntualmente los compromisos financieros derivados de la emisión.

La calificación de riesgo se efectuará de acuerdo a la metodología de calificación de cada entidad, entendiéndose como el conjunto de principios y criterios cualitativos y cuantitativos utilizados en el proceso, de acuerdo al tipo de calificación que se realice. Como resultado de todo el análisis se obtiene una calificación, que de acuerdo a nomenclatura especial, sirve para medir el riesgo implícito al valor o empresa evaluada.

Es importante mencionar y aclarar que la calificación de riesgo no constituye una sugerencia o recomendación para comprar, vender o mantener un determinado valor, ni un aval o garantía de una emisión o su emisor. Ésta es sólo la opinión de un especialista privado como un factor complementario para la toma de decisiones de inversión.

1.4.2. Técnicas Empleadas en la Construcción de Modelos

Para construir un modelo de riesgo de crédito se utilizan principalmente las siguientes técnicas:

1. Técnicas Econométricas.

Se utilizan las regresiones múltiples, análisis lineal y discriminante para determinar la probabilidad de incumplimiento como una variable dependiente. La varianza de la variable está dada por una serie de variables independientes que pueden estar directamente relacionadas con el acreditado (razones financieras) o variables externas al acreditado (tasas de interés, tipo de cambio, etc.).

2. Redes Neuronales.

Con los mismos datos que en las técnicas econométricas se crean modelos de decisión por medio de una red de neuronas interconectadas.

3. Modelos de Optimización.

Son usados para optimizar la relación entre el acreditado y los atributos del crédito para minimizar el incumplimiento y maximizar la utilidad de la institución financiera utilizando herramientas matemáticas de programación.

4. Sistemas de Expertos.

Explican de manera estructural el proceso que un analista experto usa para tomar una decisión de crédito mediante el establecimiento de una serie de reglas de decisión.

5. Sistemas Híbridos.

Mediante la estimación de parámetros y la elaboración de matrices de probabilidad de migración para predecir la tendencia de un crédito a migrar, se buscan relaciones causales directas de incumplimiento.

Aplicación de los Modelos

Los modelos son usados para:

1. Determinar calificaciones de créditos comerciales, las cuales son utilizadas para establecer límites de otorgamiento de crédito o límites a la cartera.
2. El desarrollo de modelos de análisis paramétrico en el otorgamiento de crédito; como en créditos de consumo (tarjetas de crédito, créditos personales, créditos de adquisición de automóviles, etc.) y créditos a pequeñas y medianas empresas.
3. Asignar un premio por riesgo a los créditos a partir de la probabilidad de pérdida y su dimensión en caso de ocurrir el incumplimiento.
4. Estimar el comportamiento futuro de una cartera bajo un escenario específico esperado e implementar medidas correctivas.
5. Escoger la mejor estrategia de cobranza o de recuperación de cartera (*workout*).

1.4.3. Clasificación de los Modelos

Los modelos para el cálculo del riesgo de crédito se pueden clasificar en cuatro categorías:

1. Modelos Tradicionales.
2. Modelos de Valor de Firma (modelos estructurales).
3. Modelos de Primer Tiempo de Paso.
4. Modelos Contingentes (Modelos de intensidad de *Default*⁷).

⁷Incumplimiento

Modelos Tradicionales

Se caracterizan por su facilidad en el cálculo y por eso son los más usados en la estimación del riesgo de crédito. Este tipo de modelos proporciona una estimación insuficiente.

Entre los principales tenemos:

1. Sistemas de calificación de crédito.

Éstos consisten en identificar los factores que determinan la probabilidad de incumplimiento y combinarlos en una calificación cuantitativa.

La calificación puede ser interpretada literalmente como la posibilidad de incumplimiento o se usa como un sistema de clasificación, al colocar a un acreditado potencial en un grupo malo o bueno, de acuerdo a su calificación.

2. Modelo VaR.

Desde que en 1993, el *BIS* anunciara su propósito de introducir un requerimiento para riesgo de mercado, comenzaron a surgir grandes avances en el desarrollo de metodologías de valor en riesgo. Más aún cuando en 1995 el *BIS* modificó su propuesta sobre riesgo de mercado y permitió a ciertos bancos el uso de sus propios modelos internos para el cálculo de su exposición al riesgo de mercado [6].

Los modelos VaR miden la pérdida máxima de valor de un activo o una pérdida determinada, durante un periodo de tiempo, con un nivel de confianza específico.

El VaR es una medida que busca consolidar el riesgo en un número, determinando lo máximo que de una inversión se puede perder en el peor de los escenarios. Un reporte del valor en riesgo diría: "De una inversión de 1'000.000, lo máximo que usted podría perder, con una confianza del 99 %, es 100.000".

La utilidad de este método radica en que no sólo mide el riesgo de una única inversión, sino también, lo hace para un portafolio de inver-

siones. Además, con este método se puede determinar el máximo valor que podría ganar con la inversión. Es como un indicador de los puntos extremos dentro de los cuales se puede mover la rentabilidad.

Es importante recordar que el VaR no es un instrumento infalible: es simplemente una herramienta basada en datos históricos y no contempla sucesos inesperados.

Los más importantes son:

a) *Creditmetrics*[12].

Fue introducido en 1997 por *JP Morgan* como un marco de referencia de VaR en la evaluación y riesgo de activos no negociables (créditos y bonos colocados en forma privada). Éste ha tenido gran acogida en el mundo de las finanzas, que lo considera un instrumento útil y fácil de entender.

Con información histórica se determina la máxima pérdida (o ganancia) posible en un período específico. Este cálculo puede hacerse de distintas formas, por lo que los resultados pueden diferir.

b) *RiskMetrics*[12]

Este método contesta a la pregunta ¿Si mañana es un mal día, cuánto perderé en activos negociables como acciones y bonos?, mientras que *CreditMetrics* a ¿ Si el año próximo es malo, cuánto perderé en mis créditos y en mi cartera de créditos ?

3. Basado en datos históricos.

Este modelo basado en la información histórica de bonos o créditos cotizados en la bolsa proporciona a cada compañía una calificación de riesgo asumiendo que compañías con igual calificación tienen la misma probabilidad de incumplimiento. Las principales compañías calificadoras son *Standard and Poors* [22], *Moody's*⁸ y *KMV* [12].

⁸Fué fundada en 1900 por John Moody. Sus primeras calificaciones fueron de más de 250 bonos de ferrocarriles en Estados Unidos en 1909.

El objetivo de calificar es impactar con los *spreads* o primas de riesgo de crédito requeridos sobre los flujos de efectivo del préstamo, y por tanto, sobre el valor de mercado del préstamo.

Entonces, si la calificación de un crédito baja, el premio requerido debería incrementarse, con lo que el valor presente del préstamo descendería y en caso de una mejoría en la calificación crediticia se tendría el efecto opuesto.

4. Los basados en datos de precios de bonos.

Por medio de este método se aproxima el valor de los títulos de deuda por medio de la información histórica sobre los eventos de incumplimiento.

De esta manera los bonos sujetos a riesgo de incumplimiento se valúan de forma que el inversionista sea compensado por la pérdida esperada de los títulos, con base en estimaciones a partir de los registros históricos de incumplimiento.

Las desventajas de este método es que se basa en dos supuestos frágiles:

- a) Usa el comportamiento pasado como buen predictor del desempeño futuro.
- b) Dar por hecho que los inversionistas son neutrales al riesgo (el *spread* sobre un bono libre de riesgo recompensa sólo por la pérdida esperada por el evento de incumplimiento) o bien que el riesgo de crédito no puede considerarse como riesgo sistemático, por lo cual no se requiere una remuneración extra (prima por riesgo).

Por ejemplo se tiene el caso de *Fons* (1994), que determinó el precio de los bonos de manera que fuesen consistentes con las pérdidas pasadas de instrumentos de análogas características.

Litterman e Iben (1991) y *Hurley y Johnson* (1996) derivan las probabilidades implícitas de incumplimiento a partir de los *spreads* sobre los bonos libres de riesgo. Los primeros hacen énfasis en que estas probabilidades no son empíricas, sino que por el contrario son probabilidades de riesgo neutral.

Modelos de Valor de Firma (modelos estructurales)

Este tipo de modelos toman en cuenta el valor de la firma, y aquí el incumplimiento se presenta cuando el valor alcanza un límite en una barrera determinada que depende del tiempo.

Estos modelos tienen el inconveniente de que la experiencia dice que los *spreads* no decrecen a cero cuando se acerca la fecha de vencimiento.

Además de no considerar que aunque esté muy cercana la fecha de vencimiento existe la posibilidad de un desastre en el mercado en el periodo restante para el vencimiento.

Dentro de estos modelos se encuentran los siguientes:

1. Modelo de *Merton*.

Uno de los primeros modelos para evaluar bonos e instrumentos similares sujetos a riesgo de incumplimiento, fué el desarrollado por *Merton* en 1974. Este modelo asume que la capacidad de pago de una firma está determinada por el valor de sus activos (V). Considera además, una empresa con un único pasivo, el cual tiene un único pago al vencimiento (K).

Merton parte del trabajo de *Black & Scholes*[14], donde la emisión de deuda puede interpretarse como la venta, por parte de los accionistas a los tenedores de deuda, de los activos de la firma, manteniendo los primeros una opción de compra (*Call*)⁹ sobre dichos activos. Ésto equivale a decir que los accionistas mantienen la propiedad de los activos de la empresa y compran a su vez una opción de venta (*Put*)¹⁰ a los tenedores de deuda, con un precio de ejercicio igual a K .

De esta manera, un bono sujeto a riesgo de incumplimiento puede valuarse como un bono libre de riesgo menos una opción de venta sobre los activos de la firma, con un precio de ejercicio igual a K .

⁹Da al poseedor el derecho, mas no la obligación para comprar un activo subyacente en una fecha próxima y a un cierto precio

¹⁰Da al poseedor el derecho, mas no la obligación de vender un activo subyacente en una fecha futura a un cierto precio

Así, el pago final del bono riesgoso, P , cuyo pago final prometido es K , es igual a:

$$P = K - \max(K - V, 0) = \min(V, K) \quad (1.1)$$

El proceso que rige el valor de los activos de la firma, bajo una medida de probabilidad de riesgo neutral, se especifica como un movimiento *Browniano*:

$$\partial V = rV\partial t + \gamma V\partial W \quad (1.2)$$

donde r es la tasa libre de riesgo y γ es la desviación estándar del valor de la firma.

Con la dinámica que regula el valor de la firma, el precio de un bono con riesgo de incumplimiento es igual al precio de un bono libre de riesgo al que se le sustrae el valor de una opción *Put*. la que puede valuarse usando la formula de *Black & Scholes* ¹¹

$$P(t, T) = P^*(t, T) - PUT(V, K, T - t) \quad (1.3)$$

donde $P(\cdot)$ es el precio del bono sujeto a riesgo de crédito, $P^*(\cdot)$ es el precio del bono libre de riesgo y $PUT(\cdot)$ es el precio de la opción *Put*.

A partir de esto, *Merton* muestra que la magnitud del riesgo de crédito se deriva de la razón de activos a deuda de la empresa. De hecho, esa razón puede interpretarse como el subyacente de la opción, y se dice distancia hasta el incumplimiento para referirse a ella.

Este modelo aunque muy simple se basa en una serie de supuestos fuertes. Entre los cuales están el asumir que el incumplimiento no puede ocurrir antes del vencimiento, o que no existen pagos intermedios ni diferentes niveles de prioridad en el portafolio de deuda de una empresa. Además, se supone que no existen costos de bancarrota, y la misma sólo puede ocurrir si el nivel de los activos cae por debajo del de los pasivos, con lo cual la quiebra por problemas de liquidez queda también excluida del análisis.

¹¹Si bien V no es un activo negociable, un derivado de V sí lo es (las acciones de la empresa)

2. Modelo de *Lenard y Toft* (1996)

Su modelo se basa en el modelo de *Merton*. Incorporan el riesgo de insolvencia al considerar el impacto del vencimiento de las deudas en el ejercicio óptimo de la opción incumplida por los accionistas.

3. Modelo *KMV*

En los últimos años las ideas de *Merton* ha ido en varias direcciones. Un ejemplo es la creación, por parte de la corporación *KMV* de San Francisco, de un modelo de predicción de incumplimientos (Modelo del Monitor de Crédito).

Este modelo predice y actualiza las predicciones de incumplimiento de todas las compañías importantes y los bancos que cotizan en la bolsa.

En este método el valor de las acciones de la firma es modelado por un movimiento *Browniano* geométrico. Aquí se considera la probabilidad de no poder cubrir las obligaciones con una distribución normal y el cuantil llamado distancia al incumplimiento y para éste se hace uso de datos empíricos.

Modelos de Primer Tiempo de Paso

El objetivo de estos modelos es solucionar la carencia principal del modelo de *Merton*, respecto a la posibilidad de incumplimiento previo al vencimiento del instrumento. Para esto se asume que la quiebra se produce cuando el valor de la firma cruza cierto límite predeterminado, que en general depende del tiempo (frontera de incumplimiento).

Es importante que para modelar la tasa de recuperación, se toma como una función dependiente del tiempo dado que la frontera se especifica como una función determinística del tiempo.

Entre los modelos de este tipo se tienen:

1. *Black-Cox* (1976)

Modificando el enfoque de *Merton* en lo que se refiere a modelar ciertas cláusulas gatillo (*Safety covenants*). Las cuales permiten a los tenedores de bonos forzar la quiebra de la empresa si se verifican ciertas circunstancias.

De esta forma, se intenta proteger a los inversionistas de mayores caídas en el valor de los activos de la firma. Además de restringir la capacidad de los accionistas para transferir riqueza desde los tenedores de deuda a ellos mismos, por medio del incremento en la volatilidad del valor de la empresa.

Además, permite considerar diferentes categorías de deuda en función del orden de prioridad para el cobro, ante un evento de quiebra. En este ámbito, se define como deuda *senior* a aquella que debe ser pagada antes que ningún otro acreedor reciba cualquier pago en un evento de quiebra o liquidación. En consecuencia la deuda *junior* sólo se cobrará si los tenedores *senior* han cobrado el total de la misma.

Considera que una vez que el valor de los títulos de una compañía caen por debajo de una cota, ésta se verá obligada a reestructurarse o declararse en quiebra. La cota que se usa tiene una distribución exponencial con parámetros K y λ , variables dependientes del tiempo y la función es:

$$V^d(t) = Ke^{-\lambda(T-t)}$$

2. *Box y Cox* (1976)

Este modelo encuentra dificultades para generar los premios por incumplimiento que se observan en los mercados, a menos que se supongan costos de bancarrota inusualmente altos.

3. *Brenan y Schwartz*(1980)

Presentan un modelo muy similar al de *Black y Cox*, aunque aquí la cota de incumplimiento se considera constante en la valuación para bonos convertibles.

4. *Mason y Bhattacharya* (1981)

Este modelo se vincula con un proceso de saltos. El cual contiene un tiempo de incumplimiento aleatorio para modelar el valor de la firma, considerando al tiempo de salto aleatorio y con esto logran demostrar que estos efectos tienen influencia en el precio de los bonos riesgosos.

5. *Frank y Torous (1989)*

En su modelo señalan que en muchas liquidaciones de empresas, los acreedores subordinados (incluyendo los accionistas) obtienen una porción mayor del valor residual que la consistente con respecto de la prioridad absoluta implícita en la estructura *senior/junior*.

6. *Nielsen, Saa Requejo y Santa Clara*

En este modelo la aproximación del primer tiempo de paso tiene una cota estocástica, y el incumplimiento se presenta si ésta es tocada. Se usa el modelo de *Vasicek*¹² con parámetros de variación de tiempo como *Hull y White*¹³ para estimar la tasa de recuperación.

7. *Kim, Ramaswamy y Sundarson (1993)*

En este modelo se introduce una tasa de interés estocástica. Con el proceso de raíz cuadrada de *Cox, Ingersoll y Ross (1985)* suponen que la cota de incumplimiento es una constante dependiente del tiempo.

$$V^d(t) = K$$

Además, el nivel de la cota depende del valor del cupón que se debe a los acreedores del bono y en caso de no ser tocada durante la vigencia, el pago final es de $\min(V, F)$, donde V es el valor de la firma y F el valor nominal de la deuda. En caso de que la cota sea tocada, se tiene que el incumplimiento y la tasa de recuperación es $\min(w(t)P(t, T), V)$ donde $w(t)$ es el factor del bono corporativo.

Los autores de este modelo encontraron que el *spread* de crédito es más alto al introducirse tasas de interés estocásticas y no las estándares.

¹²Desarrolló un proceso llamado regreso a la media (1977)

¹³Obtuvieron la solución acotada para el bono cupón cero, si se tiene un comportamiento *Vasicek*(1990).

8. *Longstaff-Schwartz (1995)*

Incorporan las violaciones de *Frank y Torous (1989)* a la prioridad absoluta en un modelo estocástico, donde consideran la distribución del valor residual de la firma en bancarrota entre los acreedores como dato exógeno.

Con este modelo logran mostrar que la expectativa de alejamiento respecto a la prioridad absoluta, previo a la bancarrota, pueden generar niveles de premio por riesgo similares a los observados.

Entonces, bajo este marco, el precio de un bono cupón cero sujeto a riesgo de crédito, puede escribirse como:

$$P(t, T) = P^*(t, T)(1 - \delta Q(X, r, t, T))$$

donde:

$P^*(t, T)$ es el precio del bono libre de riesgo.

r es la tasa de interés de corto plazo.

δ es el porcentaje de pérdida del valor en caso de incumplimiento.

X es la razón entre el valor de la firma V y la cota de incumplimiento, k .

Q representa la medida de probabilidad riesgo neutral de que el incumplimiento ocurra.

9. *Anderson y Sundaresan (1996)*

Incluyen un enfoque alternativo que toma como punto de partida la observación que, en muchos casos, los accionistas convencen a los tenedores de deuda de aceptar concesiones previas a la declaración formal de la bancarrota. Ésto lo presentan en un modelo binomial y examinan el diseño de contratos de deuda.

10. *Mella-Barral y Perraudin (1997)*

Derivan un modelo de tiempo continuo de valuación de deuda con servicio estratégico de la misma.

11. *Brys y Varene (1997)*

En este modelo se supone que tanto la cota de tiempo de paso y la tasa de recuperación están dados, por lo que el valor de la firma no es considerada en los cálculos.

La cota de incumplimiento es :

$$v(t) = kFP(t, T)$$

donde:

k es una constante exógena.

F es el valor nominal del bono.

$P(t, T)$ es el bono libre de riesgo.

Además, este modelo contempla el incumplimiento antes y en la fecha de vencimiento. Si éste ocurre en el vencimiento, el valor de los activos de la firma es menor que sus reponsabilidades y entonces una fracción v_1 del valor de los activos se pagará y si la cota de incumplimiento es tocada antes del vencimiento se pagará la fracción v_2 del valor de la cota $v(t)$. Las fracciones v_1 y v_2 son valores exógenos.

Modelos Contingentes (Modelos de Intensidad de *Default*)

Estos modelos representan un enfoque alternativo a los modelos estructurales. En este tipo de modelos se define un proceso estocástico que gobierna la eventualidad del incumplimiento.

En ellos, en general no se considera la jerarquía del pasivo para determinar la tasa de recuperación, por lo que, a diferencia de algunos modelos estructurales, dicha tasa es usualmente una variable exógena.

En estos modelos se demanda menor cantidad de información que en los estructurales, con lo cual se permite que se generalice su uso para valorar instrumentos con riesgo de crédito. Estos modelos valúan un bono sujeto a riesgo de incumplimiento descontando el flujo cometido a tasas mayores a las libres de riesgo.

Por ejemplo, partiendo de un conjunto de tasas *forward*¹⁴, f_t , un bono cupón

¹⁴Contratos por adelantado.

zero emitido por un emisor particular tendrá un precio

$$P = E_t(e^{\int_t^T f_r dr})$$

con $f_t > f_t^*$ y donde f_t^* representa las tasas *forward* de un emisor libre de riesgo. La diferencia entre ambas tasas representa el riesgo instantáneo de incumplimiento:

$$\lambda = f_t - f_t^*$$

donde f_t^* puede interpretarse como la **compensación** por la preferencia temporal (premio por esperar), mientras que λ_t representa la compensación por riesgo de ocurrencia de *default* entre los momentos t y $t + dt$.

Podría pensarse que el incumplimiento se dispara a partir del primer salto en un proceso de *Poisson*, con tasa λ , y que en caso de ocurrencia, la pérdida es total.

La probabilidad de n saltos para un proceso *Poisson*¹⁵ en el periodo $[T, t]$ con intensidad λ es:

$$P[n \text{ saltos}] = \frac{(\lambda[T-t])^n}{n!} e^{-\lambda(T-t)}$$

y entonces:

$$P[0 \text{ saltos}] = \frac{(\lambda[T-t])^0}{0!} e^{-\lambda(T-t)} = e^{-\lambda(T-t)}$$

Entonces, de esta forma se tiene que el precio de un bono cupón cero es:

$$P = E_t[e^{-\int_t^T f_r^* dr} (1 * e^{-\int_t^T \lambda_r dr} + 0 * (1 - e^{-\int_t^T \lambda_r dr}))] = E_t[e^{-\int_t^T (f_r^* + \lambda_r) dr}]$$

El inconveniente con este tipo de modelos es que resulta difícil relacionar este enfoque con la situación particular de un emisor.

Entre los principales modelos de este tipo tenemos los siguientes:

¹⁵Una familia de variable aleatorias N_t , relacionadas con una variable aleatoria t con dominio $[0, \infty)$ talque: $N(0) = 0$, tiene incrementos independientes y estacionarios y el número de eventos en un intervalo t es *Poisson* con media λt .

1. *Jarrow y Turnbull (1995)*

Este modelo es el primero de este tipo en tiempo continuo. Es un modelo para una economía donde el momento en el cual el incumplimiento puede ocurrir, τ , se distribuye exponencialmente con parámetro λ (denominada intensidad de incumplimiento).

Además, tanto λ como R (tasa de recuperación) son constantes, es decir, ambas tasas son independientes de cualquier variable de estado, como por ejemplo la tasa de interés.

En este modelo, el precio de un bono sujeto a riesgo de incumplimiento puede interpretarse usando una analogía de tipo de cambio externo con tipo de cambio "1" si no hay incumplimiento y $R < 1$ si éste ocurre.

En una economía libre de riesgo de incumplimiento como la del modelo *Heath-Jarrow-Morton*, el modelo admite una medida de martingala equivalente, si existe un único vector de precios de riesgo que soluciona la siguiente ecuación:

$$\alpha^*(t, T) = -b(t, T)\gamma^*(t) \quad (1.4)$$

donde $\alpha^*(t, T)$ representa la tendencia del proceso relativo bajo dicha medida y $b(t, T)$ su volatilidad. Esta ecuación muestra que la tendencia y la volatilidad son proporcionales con un factor $\gamma^*(t)$ y si es único, entonces no hay oportunidad de arbitraje.

Por otra parte, si se tiene una economía sujeta a riesgo de incumplimiento se debe de cumplir una restricción de arbitraje adicional $\forall t < \tau$. Si la tendencia del proceso es una función $\alpha^{incum} \equiv \alpha(t, T)$, entonces, se tiene que dicha condición es:

$$\alpha(t, T) = (1 - RH(t, T))\lambda\gamma(t) \quad (1.5)$$

donde $\gamma(t)$ es el precio del riesgo de incumplimiento y $H(t, T)$ es el factor de corrección de la tendencia en tasas forward, debido al evento de incumplimiento.

Y para la tasa de corto plazo, con $r^{incum}(t) \equiv r(t)$ tenemos:

$$t(t) - \tau^*(t) = (1 - R)(\lambda\gamma(t)) \quad (1.6)$$

Entonces si estas tres ecuaciones admiten una **única** solución para el par $[\tau(t), \tau^*(t)]$, esta economía está libre de **oportunidad** de arbitraje.

En este modelo se supone que $\gamma(t) = \gamma$ es **constante**, entonces, en este caso como τ se distribuye exponencial con **parámetro** λ bajo la medida empírica, se distribuye también exponencial la medida de riesgo neutral, pero con **parámetro** λ_r . Y la **probabilidad** de sobrevivir hasta $T > t$ es :

$$Q(\tau > T \mid \tau > t) = e^{-\lambda_r\gamma(T-t)} \quad (1.7)$$

Y si el mercado admite una **única** martingala se tiene que el precio del bono sujeto a riesgo de crédito es:

$$P(t, T) = B^*(t)E_Q[B^{*-1}(T)e(T) \mid F_t] \quad (1.8)$$

donde

$$e(t) = \begin{cases} 1 & \text{si } T < \tau \\ R & \text{si } T \geq \tau \end{cases}$$

$b^*(t)$ representa las letras de tesorería del emisor libre de riesgo. F_t es una filtración definida en el espacio de **probabilidad**.

2. Jarrow, Landon y Turnbull (1997)

Continuando con el trabajo de *Jarrow y Turnbull* (1995), eliminan el supuesto de que la intensidad de incumplimiento es constante todo el tiempo, pero conservando el supuesto de **independencia** y que la tasa de recuperación es constante y exógena.

Usando cadenas de *Markov* con espacio de **estados** finitos de tiempos homogéneos con una matriz generadora y la **medida** de martingala, relacionan la probabilidad de incumplimiento con las **clasificaciones** de crédito.

En la matriz generadora de transición, cada **entrada** representa a la probabilidad de cambio de clasificación, cada **entrada** se supone exponencial con **parámetro** correspondiente a la **probabilidad** de transición y la tasa de recuperación es estimada de los **datos** históricos.

Un problema en este modelo, es que en la matriz de transición se supone que las compañías con la misma clasificación tienen el mismo *spread* y que el *spread* de crédito sólo se altera cuando hay cambio de clasificación.

3. Jarrow y Turnbull (1998)

En este modelo es usado el proceso *Vasicek*, pero ignorando los valores negativos, y una función de riesgo lineal con movimiento *Browniano*. Aquí se tiene una gran flexibilidad ya que se tiene la libertad de escoger los tres parámetros y se involucra tanto al movimiento *Browniano* como la tasa *spot*, además el riesgo tiene dependencia con fundamentos económicos.

4. Duffie y Singleton

Este modelo es muy parecido al de *Jarrow y Turnbull (1995)*, pero en éste se supone que en el momento de incumplimiento (τ), el bono corporativo pierde un tanto por ciento de su valor denotado por l_τ .

Este modelo incluye un proceso tasa-incumplimiento y se puede adecuar un *spread* al rendimiento.

Si el incumplimiento pasa en un tiempo $t < \tau$, entonces el valor del bono empieza a decrecer con probabilidad l_t y con probabilidad $1 - l_t$ el bono no cambia.

Son considerados dos tipos de incumplimiento, el inofensivo con tasa de densidad $h_t(1 - l_t)$ y el peligroso con tasa de intensidad $h_t l_t$.

5. Modelo de Bonos

En este modelo el riesgo de crédito es calculado considerando las dos siguientes opciones:

- a) Si el proceso de recuperación es δ constante.
- b) Si el proceso de recuperación es cero.

Capítulo 2

Métodos de Entropía

2.1. Introducción

En muchos casos en los procesos de determinación de información nos encontramos restringidos, debido a que, el modelo fundamental muestral es incompleto o está mal especificado. Además, los datos con los que generalmente se dispone en la práctica son limitados, parciales o incompletos. Entonces, con estas restricciones, alcanzar un modelo económico-estadístico manejable puede no ser posible y los métodos convencionales pueden fallar al determinar una única solución.

Bajo estas circunstancias, definiremos el problema como mal planteado o indeterminado. Según *Jaynes*(1984), cuando se tiene un problema mal planteado o indeterminado, haciendo suposiciones y usando información anterior se puede llegar a un problema bien planteado. Este problema puede ser bien manejado con las herramientas econométricas tradicionales(modelos estadísticos y matemáticos).

El punto aquí es que deberíamos poder resolver el problema con la información incompleta o parcial y el problema mal planteado. Es importante entonces buscar bases para un razonamiento lógico en este tipo de situaciones. Además, en este tipo de casos, lo que se busca es tener un principio o formalismo con el cual podamos obtener la mejor conclusión posible basandonos en la información con que contamos. Con esta necesidad tenemos dos requerimientos esenciales:

1. Contamos con algo de información pero no con toda o quizás no con la información suficiente para proceder con un método tradicional.
2. No deseamos ni más ni menos información de la que tenemos.

Basados en todo lo anterior en la siguiente sección se presenta la solución que *Jaynes* propuso para este problema.

2.2. Máxima Entropía

2.2.1. El Problema

Considerando el problema de inversión pura finito, discreto y lineal

$$y = X\beta = Xp \quad (2.1)$$

donde, los datos están dados en forma de un vector y de dimensión T y el operador lineal X es una matriz de $(T \times (K > T))$ no invertible, pero conocida. Deseamos determinar las desconocidas y no observables frecuencias $p = (p_1, \dots, p_K)^t$ que representan el proceso de generación de datos. Entonces, nuestro objetivo es recuperar, asignar o escoger probabilidades no ambiguas p_k de entre todas las distribuciones de probabilidad que satisfagan (2.1).

Pero dadas estas especificaciones, la información contenida en (2.1) no parece la adecuada para determinar las probabilidades desconocidas. Ésto se debe a que el número de datos es menor que el número de incógnitas. Hasta aquí el problema sigue pareciendo indeterminado, al igual que las bases para asignar una probabilidad.

Dada la información limitada, si deseamos hacer una inferencia de las probabilidades, entonces, deberíamos escoger alguna estimación

$$\hat{p} = Ay \quad (2.2)$$

donde, A es un operador desconocido para ser escogido. Aquí la cuestión es como escoger a A . Si en la selección usamos las leyes de la lógica, entonces la A seleccionada debería producir \hat{p} en el conjunto de posibles $\{p\}$ que satisfacen

$$y = X\hat{p} \implies XAy = XAXp \quad (2.3)$$

entonces dada (2.3) tenemos que

$$XAX = X \quad (2.4)$$

entonces A es inversa generalizada (Kalman, 1960)[3].

Esta restricción nos ayuda en el sentido de que nos permite identificar el conjunto de soluciones en el cual la verdadera p se encuentra. Pero desafortunadamente, ésto no proporciona las bases para hacer una selección de A o p en este conjunto, ya que \hat{p} no puede distinguirse de la verdadera p .

2.2.2. Medida de Entropía

Requisitos de una medida de incertidumbre de una distribución de probabilidad

Sean p_1, \dots, p_K las probabilidades de K posibles resultados N_1, \dots, N_K de un experimento, dando como resultado la distribución de probabilidad

$$p = (p_1, \dots, p_K)^t$$

sujeto a:

$$\sum_{k=1}^K p_k = 1$$

y

$$p_k \geq 0$$

pero aquí hay una incertidumbre cuando el experimento es realizado.

Una medida de esta incertidumbre debería de satisfacer los siguientes requisitos:

1. Debe ser función de p_1, \dots, p_K de manera que se pudiera escribir como

$$H = H_K(p) = H_K(p_1, \dots, p_K)$$

2. Ser una función continua de p_1, \dots, p_K , en el sentido de que, un cambio pequeño en p_1, \dots, p_K debería causar un cambio pequeño en H_K .

3. No debe cambiar bajo un reordenamiento de los resultados, es decir, H_K es una función simétrica de éstos.
4. No debe cambiar si un resultado imposible es añadido al esquema de probabilidad, es decir,

$$H_{K+1}(p_1, \dots, p_K, 0) = H_K(p_1, \dots, p_K)$$

5. Debe ser mínima y posiblemente cero en caso de que no exista incertidumbre acerca de los resultados. Esto debería presentarse cuando uno de los resultados es seguro y entonces

$$H_K(p_1, \dots, p_K) = 0$$

cuando $p_i = 1, p_j = 0$ para $i \neq j$ y $i = 1, \dots, K$.

6. Debe ser máxima cuando hay máxima incertidumbre, la cual se tiene cuando los resultados son equiprobables, es decir, H_K es máxima cuando $p_1 = \dots = p_K = \frac{1}{K}$.
7. El valor máximo de H_K aumenta cuando K aumenta.
8. Para dos distribuciones de probabilidad independientes $p = (p_1, \dots, p_K)^t$ y $q = (q_1, \dots, q_M)^t$, donde si N_1, \dots, N_K y B_1, \dots, B_M son los resultados de p y q respectivamente, entonces los resultados de $p \cup q$ son $N_i B_j$ con probabilidades $p_i q_j$ para $i = 1, \dots, K$ y $j = 1, \dots, M$, la incertidumbre de $p \cup q$ debe ser la suma de las incertidumbres

$$H_{K+M}(p \cup q) = H_K(p) + H_M(q).$$

La Medida de Shannon

Bajo situaciones como la descrita en la introducción de este capítulo, *Shannon* (1948), quien quería bases para medir la incertidumbre en algunas mentes al recibir un mensaje ruidoso, utilizó un método axiomático para definir una única función para medir la incertidumbre de la ocurrencia de una

colección de eventos.

Sea X una variable aleatoria con posibles valores x_k , $k = 1, \dots, K$ y p_k sus respectivas probabilidades, tales que, $\sum_k p_k = 1$. *Shannon* definió la entropía de la distribución de probabilidades $p = (p_1, \dots, p_K)^t$ como la medida

$$H(p) = - \sum_k p_k \ln(p_k) \quad (2.5)$$

la cual, definida de esta forma, es fácil ver que es función de p_1, \dots, p_K . Ésta también es una función continua y es simétrica siempre que $0 * \ln(0) = 0$. Ésta no cambia cuando un resultado imposible se añade al esquema de probabilidad. Cuando una de las probabilidades es uno y las otras son cero, el valor de H_p es cero y éste es el valor mínimo que alcanza, pues $H(p) \geq 0$ cuando $0 \leq p_k \leq 1$.

Por otro lado, como $x \ln(x)$ es una función convexa, entonces $\sum_k p_k \ln(p_k)$ es una función convexa, y entonces $-\sum_k p_k \ln(p_k)$ es una función concava y el máximo local es un máximo absoluto. Éste se alcanza cuando $p_1 = p_2 = \dots = \frac{1}{K}$ (usando multiplicadores de *Lagrange*), es decir, cuando las probabilidades son uniformes.

El valor máximo de $H(p)$ es

$$- \sum_{k=1}^K \frac{1}{K} \ln\left(\frac{1}{K}\right) = \ln(K)$$

Otra manera de ver que $H(p)$ se maximiza cuando $p_1 = p_2 = \dots = \frac{1}{K}$, es usando la desigualdad de *Jensen* para una función convexa $f(x)$, dicha desigualdad es:

$$E[f(x)] \geq f[E(x)]$$

para alguna variable aleatoria x . Tomemos $f(x) = x \ln x$ y que x tome los valores p_1, \dots, p_K , cada uno con probabilidad $\frac{1}{K}$, entonces

$$E[f(x)] = \frac{1}{K} \sum_k p_k \ln(p_k)$$

mientras que

$$E(x) = \frac{1}{K} \sum_k p_k = \frac{1}{K}$$

entonces

$$f[E(x)] = f\left(\frac{1}{K}\right) = \frac{1}{K} \ln\left(\frac{1}{K}\right)$$

usando la desigualdad de *Jensen* tenemos

$$\sum_k p_k \ln(p_k) \geq \ln\left(\frac{1}{K}\right)$$

o

$$-\sum_k p_k \ln(p_k) \leq \ln(K) = -\sum_{k=1}^K \frac{1}{K} \ln\left(\frac{1}{K}\right).$$

Principio de Máxima Entropía de Jaynes

El concepto de entropía, se remonta años atrás con *Boltzman* (en los 1870s), así como *Maxwell*, *Gibbs*, *Shannon* y el trabajo realizado por *Bernoulli*, *Laplace*, *Jeffreys* y *Cox*; *Jaynes* (1957) propuso hacer uso de la entropía para seleccionar la desconocida distribución de la probabilidad de (2.1).

Jaynes, llamó principio de máxima entropía a la selección de la distribución de probabilidad, para la cual la información que se tiene, es la suficiente para determinar la probabilidad, es decir, maximizar la entropía de *Shannon* sujeta a las restricciones .

La propuesta se desarrolla realizando N pruebas (repeticiones) de un experimento que tiene K posibles resultados (estados). Sean N_1, \dots, N_K el número de veces que cada resultado ocurre en el experimento de N repeticiones, donde

$$\sum_{k=1}^K N_k = N \quad (2.6)$$

y $N_k \geq 0$.

Ya que son realizadas N pruebas y cada una tiene K posibles resultados,

entonces hay K^N posibles resultados en la sucesión de N pruebas. De éstas un conjunto particular de sucesiones

$$p_k = \frac{N_k}{N} \quad (2.7)$$

o bien $N_k = p_k N$ para $k = 1, \dots, K$, puede ser realizado en un número dado de formas medidas por el factor de multiplicidad (posibles permutaciones).

Entonces, podemos representar el número de formas en que un conjunto N_k puede ser realizado por el coeficiente multinomial

$$W = \frac{N!}{N p_1! \cdot \dots \cdot N p_k!} = \frac{N!}{\prod_k N_k!} \quad (2.8)$$

o la función monótona de W

$$\ln(W) = \ln(N!) - \sum_{k=1}^K \ln(N_k!) \quad (2.9)$$

Con la aproximación de *Stirling*

$$\ln(x!) \approx x \ln(x) - x \quad (2.10)$$

cuando $0 < x \rightarrow \infty$, para aproximar cada componente del lado derecho de la ecuación (2.9). Tenemos, para N muy grande

$$\ln(W) \approx N \ln(N) - N - \sum_{k=1}^K N_k \ln(N_k) + \sum_{k=1}^K N_k \quad (2.11)$$

sustituyendo (2.6), tenemos

$$\ln(W) \approx N \ln(N) - \sum_{k=1}^K N_k \ln(N_k) \quad (2.12)$$

La razón $\frac{N_k}{N}$ representa la frecuencia de ocurrencia de los posibles K resultados en la sucesión de tamaño N y

$$\frac{N_k}{N} \rightarrow p_k$$

cuando $N \rightarrow \infty$.

De (2.12) tenemos que

$$\ln(W) \approx N \ln(N) - \sum_{k=1}^K N p_k \ln(N p_k) \quad (2.13)$$

o

$$\ln(W) \approx N \ln(N) - \sum_{k=1}^K N_k \ln(N) - N \sum_{k=1}^K p_k \ln(p_k) \quad (2.14)$$

o

$$\ln(W) \approx -N \sum_{k=1}^K p_k \ln(p_k) \quad (2.15)$$

finalmente

$$N^{-1} \ln(W) \approx - \sum_{k=1}^K p_k \ln(p_k) = H(p) \quad (2.16)$$

que es la medida de entropía de *Shannon* y donde $p_k \ln(p_k) = 0$ cuando $p_k = 0$.

La entropía de (2.5) o (2.16) es maximizada con el valor máximo de $\ln(K)$, cuando $p_1 = p_2 = \dots = \frac{1}{K}$.

Dada (2.16), siguiendo a *Jaynes* y maximizando la función monótona de W sujeta a las restricciones y los datos de (2.1) obtenemos un conjunto p_k (distribuciones de frecuencia) que pueden ser realizadas con el mayor número de datos consistentes con que contamos.

Entonces H es una medida de la incertidumbre en la distribución de probabilidad y es numericamente igual a la distribución de probabilidad que maximiza la entropía de *Shannon-Jaynes* donde

$$W = e^{NH}. \quad (2.17)$$

Así, si queremos saber cual conjunto particular de frecuencias **relativa** es la mejor aproximación para p_k , es razonable seguir a *Jaynes* y **favorecer a la** que fué generada con el mayor número de datos consistentes con que contamos.

Entonces, debemos escoger la p que maximize

$$H(p) = - \sum_k p_k \ln(p_k) = -p^t \ln(p) \quad (2.18)$$

donde $\ln(p)$ es un vector de $K \times 1$ y sujeta a las restricciones de consistencia de datos y normalización de la adición

$$y = Xp \quad (2.19)$$

y

$$p^t 1 = 1 \quad (2.20)$$

donde 1 es un vector de unos de $K \times 1$

Ahora, como consecuencia de ésto, tenemos que convertir nuestro problema de matemáticas deductivas a uno de inferencia. En éste buscamos hacer la mejor predicción posible de p con la información que contamos.

Por lo general, la información (los datos) pueden ser representados por T funciones $\{f_1(x), f_2(x), \dots, f_T(x)\}$. Bajo el formalismo de la máxima entropía, el problema de recuperar p está dado por

$$\max H(p) = - \sum_{k=1}^K p_k \ln(p_k) \quad (2.21)$$

sujeto a la restricción de consistencia de momento

$$\sum_{k=1}^K p_k f_t(x_k) = y_t \quad (2.22)$$

con $1 \leq t \leq T$.

Y la restricción de normalización aditiva

$$\sum_{k=1}^K p_k = 1 \quad (2.23)$$

donde $\{y_1, \dots, y_T\}$ es el conjunto de datos observados que son consistentes con

la distribución de probabilidad $\{p_1, \dots, p_K\}$.

Es importante mencionar que el problema está mal planteado o indeterminado siempre que $T < K$.

Ahora, para obtener el vector de probabilidad p se utiliza la función de *Lagrange*

$$L = - \sum_{k=1}^T p_k \ln(p_k) + \sum_{t=1}^T \lambda_t [y_t - \sum_{k=1}^K p_k f_t(x_k)] + \mu (1 - \sum_{k=1}^K p_k) \quad (2.24)$$

con las condiciones de optimización de primer orden tenemos

$$\frac{\partial L}{\partial p_k} = -\ln \hat{p}_k - 1 - \sum_{t=1}^T \hat{\lambda}_t f_t(x_k) - \hat{\mu} = 0 \quad (2.25)$$

con $k = 1, 2, \dots, K$

$$\frac{\partial L}{\partial \lambda_t} = y_t - \sum_{k=1}^K \hat{p}_k f_t(x_k) = 0 \quad (2.26)$$

con $t = 1, 2, \dots, T$

$$\frac{\partial L}{\partial \mu} = 1 - \sum_{k=1}^K \hat{p}_k = 0. \quad (2.27)$$

Aquí tenemos un sistema de $K + T + 1$ ecuaciones y parámetros dados, cuya solución es:

de (2.25) obtenemos

$$\hat{p}_k = \exp\left(- \sum_{t=1}^T \hat{\lambda}_t f_t(x_k) - 1 - \hat{\mu}\right) \quad (2.28)$$

con $k = 1, 2, \dots, K$

de (2.26) tenemos

$$\sum_{k=1}^K \exp\left(- \sum_{t=1}^T \hat{\lambda}_t f_t(x_k) - 1 - \hat{\mu}\right) f_t(x_k) = y_t \quad (2.29)$$

con $t = 1, 2, \dots, T$
y de (2.27)

$$\sum_{k=1}^K \exp\left(-\sum_{t=1}^T \hat{\lambda}_t f_t(x_k) - 1 - \hat{\mu}\right) = 1. \quad (2.30)$$

La solución formal es

$$\hat{p}_k = \frac{1}{\Omega(\hat{\lambda}_1, \dots, \hat{\lambda}_T)} \exp\left(-\sum_{t=1}^T \hat{\lambda}_t f_t(x_k)\right) \quad (2.31)$$

donde

$$\Omega(\hat{\lambda}) = \sum_{k=1}^K \exp\left(-\sum_{t=1}^T \hat{\lambda}_t f_t(x_k)\right)$$

es un factor de normalización que convierte las probabilidades relativas en probabilidades absolutas.

Las $\{\hat{\lambda}_t\}$ son multiplicadores de *Lagrange* en las restricciones (2.22) y son determinados por las T ecuaciones simultaneas

$$y_t = \frac{\partial}{\partial \hat{\lambda}_t} \ln(\Omega) \quad (2.32)$$

con $1 \leq t \leq T$.

El valor de la medida de entropía H es una función de estos datos:

$$H = \ln(\Omega(\hat{\lambda})) + \sum_t \hat{\lambda}_t y_t. \quad (2.33)$$

2.2.3. Entropía Normalizada

Dada la información en un conjunto de datos, si lo que queremos medir es la importancia de la contribución de cada dato o restricción en la reducción de incertidumbre, entonces, se define la medida de entropía normalizada (Golan, 1988; Soofi, 1992).

En la máxima entropía, el nivel máximo de incertidumbre resulta cuando las restricciones de consistencia de momento (2.22) no son impuestas y la distribución de probabilidad de los K estados es uniforme.

Como cada que se agrega un dato a la información, la correspondiente distribución uniforme implica una reducción en la incertidumbre, entonces, estamos interesados en saber que proporción de incertidumbre se reduce cada que se agrega un dato a la información que ya se tiene. La proporción del total de incertidumbre que se va quedando es medida por la entropía normalizada

$$S(\hat{p}) = \frac{(-\sum_k \hat{p}_k \ln \hat{p}_k)}{\ln(K)} \quad (2.34)$$

donde $S(\hat{p}) \in [0, 1]$ y $\ln(K)$ representa la máxima incertidumbre. El valor $S(\hat{p}) = 0$ implica que no hay incertidumbre, es decir, $p_k = 1$ para algún k y $p_i = 0$ para todo $i \neq k$. Por otra parte, si $S(\hat{p}) = 1$ implica incertidumbre perfecta, es decir, $p_k = \frac{1}{K}$ para todo $k = 1, \dots, K$.

Una medida análoga $1 - S(\hat{p})$, llamada el índice de información, es usado para medir la reducción en la incertidumbre (Soofi, 1992).

Dado que $S(\hat{p})$ es una medida relativa de incertidumbre, la podemos usar para comparar diferentes casos o escenarios.

Si dado un escenario determinado, podemos agregar o quitar un dato de la información o una restricción y comparar $S(T)$ con $S(T - 1)$, entonces tenemos los siguientes posibles casos:

1. Si $S(T) = S(T - 1)$, se puede concluir que no hay información adicional en el dato o restricción que fue agregado o retirado.
2. Si $S(T) < S(T - 1)$, entonces, el dato adicional da más información y reduce la incertidumbre sobre el conjunto de las desconocidas probabilidades.

En otras palabras, sea $\hat{p}_0(x)$ la distribución de la máxima entropía para los T datos (restricciones), y $\hat{p}_1(x)$ la distribución de la máxima entropía para los $T - 1$ datos (restricciones). Se tiene que (Levine(1980))

$$0 \leq \sum_k \hat{p}_{0k} \ln \left(\frac{\hat{p}_{0k}}{\hat{p}_{1k}} \right) = H(\hat{p}_1) - H(\hat{p}_0) \quad (2.35)$$

donde son usadas las dos desigualdades de *Gibb* y el hecho de que \hat{p}_0 es consistente con todos los primeros $T - 1$ datos (restricciones) usados en la caracterización de \hat{p}_1 . La parte de la igualdad en (2.35) se da si y sólo si $\hat{p}_0 = \hat{p}_1$. Entonces, (2.35) implica que $\hat{p}_0 \leq \hat{p}_1$.

2.3. Mínima Entropía Relativa

En muchas ocasiones en la práctica nos encontramos con que deberíamos tener una pre-muestra de información sobre las desconocidas $p = (p_1, \dots, p_K)^t$ en forma de una distribución de probabilidad anterior $q = (q_1, \dots, q_K)^t$. Es decir, antes de usar los datos (2.19) que entran como una relación consistente, debería haber alguna distribución q que pudiera ser tomada como una buena hipótesis inicial.

Si existe dicha distribución, deberíamos incorporarla a la máxima entropía.

De acuerdo con *Kullback*(1959) y *Good*(1963), el principio de mínima entropía relativa implica una selección, dadas las condiciones de (2.19) y (2.20), de la estimación de p que puede ser distinguida de q con una mínima diferencia.

Con estas primicias para la entropía relativa entre p y q , ésta se define como

$$I(p, q) = \sum_{k=1}^K p_k \ln\left(\frac{p_k}{q_k}\right) = \sum_k p_k \ln(p_k) - \sum_k p_k \ln(q_k) = p^t \ln p - p^t \ln q \quad (2.36)$$

o la diferencia entre $E[\ln(p)]$ y $E[\ln(q)]$.

Este criterio conduce a una medida natural de la desviación de las distribuciones de probabilidad p y q . Bajo el principio de mínima entropía relativa la diferencia $I(p, q)$ es minimizada. Tomando en cuenta lo anterior y la información con la que se cuenta, la solución de la mínima entropía relativa debería ser obtenida de la minimización del problema

$$\min I(p, q) = \sum_{k=1}^K p_k \ln\left(\frac{p_k}{q_k}\right) = \sum_k p_k \ln(p_k) - \sum_k p_k \ln(q_k) = p^t \ln p - p^t \ln q \quad (2.37)$$

sujeto a la restricción de consistencia de momento

$$\sum_{k=1}^K p_k f_t(x_k) = y_t$$

con $1 \leq t \leq T$.

Y la restricción de normalización aditiva

$$\sum_{k=1}^K p_k = 1.$$

Si q es consistente con los datos, entonces $\bar{p} = q$ y (2.37) tiene una solución cero, lo cual quiere decir que los datos no tienen información adicional acerca de la q .

Si con una entropía relativa, q es uniforme, resulta entonces la solución de máxima entropía (2.35). Además, el criterio de mínima entropía y la minimización del problema (2.37), así como los datos y las probabilidades anteriores son transformados en la estimación posterior de probabilidad \bar{p} .

Ahora, la probabilidad del vector p puede ser calculado con la función de *Lagrange*:

$$L = \sum_{k=1}^K p_k \ln\left(\frac{p_k}{q_k}\right) + \sum_{t=1}^T \lambda_t \left[y_t - \sum_{k=1}^K p_k f_t(x_k) \right] + \mu \left(1 - \sum_{k=1}^K p_k \right) \quad (2.38)$$

con las condiciones de optimización $\frac{\partial L}{\partial (\cdot)} = 0$.

Obtenemos el sistema de $T + K + 1$ ecuaciones y parámetros \bar{p} , $\bar{\lambda}$ y $\bar{\mu}$.

La solución, en el marco tradicional de la máxima entropía, que combina la información de los datos y la de la q anterior es:

$$\bar{p}_k = \frac{q_k}{\Omega(\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_T)} \exp\left[\sum_{t=1}^T \bar{\lambda}_t f_t(x_k) \right] \quad (2.39)$$

donde $k = 1, 2, \dots, K$ y

$$\Omega(\bar{\lambda}) = \sum_{k=1}^K q_k \exp\left[\sum_{t=1}^T \bar{\lambda}_t f_t(x_t)\right]$$

es la función de partición.

La entropía normalizada $S(\hat{p})$ se generaliza para

$$S(\hat{p}) = \frac{-\sum_k \hat{p}_k \ln(\hat{p}_k)}{-\sum_k q_k \ln(q_k)}. \quad (2.40)$$

Proposición 2.3.1 $I(p, q) = \sum_k p_k \ln\left(\frac{p_k}{q_k}\right) \approx \sum_k \frac{1}{q_k} (p_k - q_k)^2$

siempre que $q_k > 0$

Demostración:

(i)

Dada el desarrollo de Taylor para $\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \dots$
 $\forall x \in (-1, 1]$

entonces $\ln(y) = (y-1) - \frac{1}{2}(y-1)^2 + \frac{1}{3}(y-1)^3 - \frac{1}{4}(y-1)^4 + \dots$
 $\forall y \in (0, 2]$.

En general $\ln(x) \approx (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 + \dots$

(ii)

Si $p_k \approx q_k$ entonces $\ln\left(\frac{p_k}{q_k}\right) \approx \left(\frac{p_k}{q_k}\right) - 1$

Entonces

$$\sum_k p_k \ln\left(\frac{p_k}{q_k}\right) \approx \sum_k p_k \left(\frac{p_k}{q_k} - 1\right) = \sum_k \left(\frac{p_k^2}{q_k} - p_k\right) = \sum_k \frac{1}{q_k} (p_k^2 - p_k q_k) \quad (2.41)$$

y como $\sum_k q_k = \sum_k p_k = 1$ reescribimos (2.41) como:

$$\sum_k \frac{1}{q_k} (p_k^2 - p_k q_k) + 1 - 1 = \sum_k \frac{1}{q_k} (p_k^2 - p_k q_k) + \sum_k q_k - \sum_k p_k$$

$$\begin{aligned}
 &= \sum_k \frac{1}{q_k} (p_k^2 - p_k q_k) + \sum_k \frac{1}{q_k} (q_k)^2 - \sum_k \frac{1}{q_k} (p_k q_k) \\
 &= \sum_k \frac{1}{q_k} (p_k^2 - 2p_k q_k + q_k^2) \\
 &= \sum_k \frac{1}{q_k} (p_k - q_k)^2 \blacksquare
 \end{aligned}$$

Aplicando la proposición (2.3.1) a la función objetivo de la máxima entropía, donde $q_k = \frac{1}{K}$ para todo $k = 1, 2, \dots, K$ se tiene que:

$$\sum_k p_k \ln\left(\frac{p_k}{q_k}\right) = \sum_k p_k \ln(K p_k) = \sum_k (p_k \ln(p_k) + \ln(K)) \approx \sum_k K \left(p_k - \frac{1}{K}\right)^2 \quad (2.42)$$

2.4. Entropía en Variables Aleatorias

2.4.1. Variables Aleatorias Discretas

Sea X una variable aleatoria discreta con posibles valores $x = 0, 1, \dots, \infty$ y sea $p(x) = P(X = x)$, la distribución de máxima entropía con media y , se obtiene maximizando

$$H(p) = \sum_{x=0}^{\infty} p(x) \ln(p(x)) \quad (2.43)$$

sujeto a

$$\sum_{x=0}^{\infty} x p(x) = y$$

y

$$\sum_{x=0}^{\infty} p(x) = 1.$$

Para obtener la solución de este problema se utiliza nuevamente una función de *Lagrange*.

En la solución del problema, si el multiplicador de *Lagrange* $\hat{\lambda} > 0$ entonces

la solución al problema es la distribución geométrica

$$p(x) = [1 - \exp(-\hat{\lambda})] \exp(-\hat{\lambda}x) \quad (2.44)$$

con parametro $\theta = 1 - \exp(-\hat{\lambda})$.

Golan (1988,1994) desarrolló una interesante extensión del resultado para el caso en el cual la variable aleatoria es $x = 0, \epsilon, 2\epsilon, \dots, \infty$ para alguna $\epsilon > 0$ y restricciones lineales. En este caso la distribución de máxima entropía puede ser escrita como

$$p(x) = (1 - Q)Q^x \quad (2.45)$$

donde $Q = \exp(-\lambda\epsilon)$. La media de la distribución es

$$b(\epsilon) = \sum_{x=0}^{\infty} x(1-Q)Q^x = (1-Q)Q \sum_{x=0}^{\infty} xQ^{x-1}$$

$$= (1-Q)Q \frac{d}{dQ} (1-Q)^{-1} = (1-Q)Q(1-Q)^{-2} = \frac{Q}{1-Q} = \frac{1}{Q^{-1}-1}$$

sustituyendo Q tenemos

$$b(\epsilon) = \frac{1}{\exp(\lambda\epsilon) - 1} \quad (2.46)$$

la cual es la distribución de *Bose-Einstein*.

En los casos donde K es finita, escribiremos el factor normalizador como

$$\Omega(\lambda) = \sum_{k=0}^K \exp(-\sum_t \lambda_t x_{tk}) - 1$$

donde 1 representa el caso donde $k = 0$.

Si tenemos que $Q_x \equiv \exp(-\sum_t \lambda_t x_{tk})$ y $Q = \prod_k Q_k$ entonces:

$$\Omega(\lambda) = -1 + \sum_{k=1}^K Q^k = -1 + \frac{1 - Q^{K+1}}{1 - Q} = \frac{Q - Q^{K+1}}{1 - Q}$$

$$= \frac{1 - Q^K}{Q^{-1} - 1} = \frac{1 - \exp(-K \sum_t \lambda_t x_{tk})}{\exp(-\sum_t \lambda_t x_{tk}) - 1}$$

para $|Q_k| < 1$
sustituyendo en p_k tenemos

$$p_k = \frac{1}{[1 - \exp(K \sum_t \lambda_t x_{tk})][\exp(\sum_t \lambda_t x_{tk}) - 1]^{-1}[\exp(\sum_t \lambda_t x_{tk})]} \quad (2.47)$$

2.4.2. Variables Aleatorias Continuas

Sea X una variable aleatoria continua con función de densidad $f(x)$, la entropía está definida como:

$$H = - \int f(x) \ln(f(x)) dx. \quad (2.48)$$

Si la única restricción es la normalización $\int f(x) dx = 1$, entonces la función de Lagrange es:

$$\begin{aligned} L &= - \int f(x) \ln(f(x)) dx + \mu (1 - \int f(x) dx) \\ &= \mu + \int [-f(x) \ln(f(x)) - \mu f(x)] dx. \end{aligned} \quad (2.49)$$

Entonces, usando el cálculo de variaciones, la condición óptima es la ecuación (Euler)[3]

$$-d[f(x) \ln(f(x)) - \mu f(x)]/dx = 0 \quad (2.50)$$

o

$$(\ln(\hat{f}(x)) + 1 + \hat{\mu}) \hat{f}(x) = 0 \quad (2.51)$$

entonces

$$\hat{f}(x) = \exp(-1 + \hat{\mu}). \quad (2.52)$$

Dado que esta función es independiente de X , ésta puede tomar valores en los reales y $f(x)$ es una densidad impropia si la restricción es ignorada. Sin embargo, acotando el espacio de los parámetros por medio de una restricción podemos obtener una solución propia.

Con lo anterior, si la variable X es restringida al intervalo (a, b) , entonces $f(x) = \frac{1}{b-a}$ es una distribución uniforme con media $\frac{a+b}{2}$ y varianza $\frac{(b-a)^2}{12}$ donde $\mu = \ln(b-a) - 1$, y el valor de la entropía es $-\ln(b-a)$.

Por otra parte, si lo que tenemos es una variable X no negativa, definida en $[0, +\infty)$ y la esperanza es conocida, es decir, sabemos que

$$\int_0^{\infty} xf(x)dx = a \quad (2.53)$$

entonces la función de Lagrange es:

$$\begin{aligned} L &= -\int_0^{\infty} f(x)\ln(f(x))dx + \mu(1 - \int_0^{\infty} f(x)dx) + \lambda(a - \int_0^{\infty} xf(x)dx) \\ &= -\lambda a + \mu - \int_0^{\infty} [f(x)\ln(f(x)) + \mu f(x) + \lambda xf(x)]dx \end{aligned} \quad (2.54)$$

usando el cálculo de variaciones, la condición óptima es la ecuación (Euler)[3]

$$-d[f(x)\ln(f(x)) + \mu f(x) + \lambda xf(x)]/dx = 0 \quad (2.55)$$

o

$$-(\ln(\hat{f}(x)) + 1 + \hat{\mu} + \hat{\lambda}x) = 0. \quad (2.56)$$

Resolviendo la ecuación junto con las restricciones para $f(x)$, tenemos que

$$\hat{f}(x) = \frac{1}{a} \exp\left(\frac{-x}{a}\right) \quad (2.57)$$

la cual es una distribución exponencial con media a y varianza a^2 . Además la entropía tiene valor de $1 + \ln(a)$.

Ahora, si tenemos las restricciones de que la media es cero y la varianza es σ^2 , es decir, tenemos que

$$\int xf(x)dx = 0 \quad (2.58)$$

y

$$\int x^2 f(x)dx = \sigma^2 \quad (2.59)$$

entonces, la correspondiente función de *Lagrange* es

$$\begin{aligned} L &= - \int f(x) \ln(f(x)) dx + \mu(1 - \int f(x) dx) \\ &\quad + \lambda_1(0 - \int x f(x) dx) + \lambda_2(\sigma^2 - \int x^2 f(x) dx) \\ &= \mu + \int [-f(x) \ln(f(x)) - \mu f(x) - \lambda_1 x f(x) - \lambda_2 x^2 f(x)] dx \end{aligned} \quad (2.60)$$

usando el cálculo de variaciones, la condición óptima es la ecuación (*Euler*)[3]

$$-\ln(\hat{f}(x)) - 1 - \hat{\mu} - \hat{\lambda}_1 x - \hat{\lambda}_2 x^2 = 0 \quad (2.61)$$

y la solución es

$$\hat{f}(x) = \exp(-1 - \hat{\mu} - \hat{\lambda}_1 x - \hat{\lambda}_2 x^2) + \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}} \exp\left(\frac{-x^2}{2\sigma^2}\right) \quad (2.62)$$

donde

$$\hat{\mu} = \frac{\ln(2\pi\sigma^2)}{2} - 1$$

y

$$\hat{\lambda}_1 = 0, \hat{\lambda}_2 = \frac{1}{2\sigma^2}$$

y la variable aleatoria x tiene una distribución de densidad normal con media cero y varianza σ^2 .

Para este caso, el valor de la máxima entropía es

$$\begin{aligned} H(\hat{p}) &= - \int f(x) \ln(f(x)) dx = \int f(x) \left[\frac{\ln(2\pi\sigma^2)}{2} + \frac{x^2}{2\sigma^2} \right] dx \\ &= \frac{\ln(2\pi\sigma^2)}{2} + \frac{1}{2} = \frac{\ln(2e\pi\sigma^2)}{2} \end{aligned} \quad (2.63)$$

Capítulo 3

Matrices de Migración de Crédito

En este capítulo se presenta una breve introducción al concepto de calificaciones crediticias y su importancia en el mundo financiero. Además se describen los principales resultados sobre matrices de transición y se exponen diferentes métodos para estimar matrices de transición crediticias.

3.1. Calificaciones Crediticias

En los contratos de crédito, las posibles pérdidas económicas involucran la estimación de las probabilidades de incumplimiento por parte de los acreditados. El acreditado es clasificado según la capacidad de cumplir con sus obligaciones en un contrato de crédito.

En el campo de incumplimiento por parte de las empresas, en un contrato de crédito, las matrices de transición representan las probabilidades de pasar de una calificación a otra en un intervalo de tiempo.

A continuación se presentan las clasificaciones de las empresas según su capacidad de cumplimiento en un contrato de crédito.

	Categorías de clasificación de valores de deuda a corto plazo
AAA	Muy alta capacidad de pago de capital e intereses, la cual no se vería afectada ante posibles cambios en el emisor, en el sector al que pertenece o en la economía.
AA	Alta capacidad de pago de capital e intereses, la cual no se vería afectada ante posibles cambios en el emisor, en el sector al que pertenece o en la economía.
A	Buena capacidad de pago de capital e intereses, la cual es susceptible a deteriorarse levemente ante posibles cambios en el emisor, en el sector al que pertenece o en la economía.
BBB	Suficiente capacidad de pago de capital e intereses, la cual es susceptible a debilitarse ante posibles cambios en el emisor, en el sector al que pertenece o en la economía.
BB	Cuenta con capacidad de pago de capital e intereses, la cual es variable y susceptible a debilitarse ante posibles cambios en el emisor, en el sector al que pertenece o en la economía, pudiendo incurrirse en retraso de pago de intereses y del capital.
B	Cuenta con una mínima capacidad de pago, siendo muy variable y susceptible ante posibles cambios en el emisor, en el sector al que pertenece o en la economía, pudiendo incurrirse en la pérdida del capital e intereses.
C	No cuenta con la capacidad para el pago, existiendo alto riesgo de pérdida de capital e intereses.
D	No cuenta con la capacidad de pago de capital e intereses, y representa incumplimiento efectivo de pago de capital e intereses.
E	Valores cuyo emisor no posee información suficiente o no tiene información representativa para el periodo mínimo exigido para la calificación y carece de garantías suficientes.

3.1.1. Compañías Calificadoras

Las clasificaciones son emitidas por las compañías calificadoras. El papel de estas compañías en los mercados globales de capital es tener una medición de riesgo de crédito de manera independiente, creíble y objetiva; una cobertura comprensible y consistencia global; una transparencia crediticia y

un aumento de la eficiencia y la liquidez.

El mercado de calificadoras de crédito está dominado por dos compañías: *Standard and Poors (S&P)*[22] y *Moody's Investor Services (Moody's)*[21]. Una calificación de crédito representa una tasa global del cumplimiento de las obligaciones por parte del acreditado.

Existen algunas diferencias entre las calificaciones de las diferentes compañías. Mientras *S&P* evalúa la capacidad financiera del acreditado para cumplir con sus obligaciones y de esta forma estimar la probabilidad de incumplimiento, *Moody's* califica incorporando algunos juicios para la recuperación en los eventos de pérdida. En el lenguaje de la administración de riesgo de crédito se dice que *S&P* mide la probabilidad de incumplimiento, mientras *Moody's* mide la distancia a la pérdida o a la esperanza de pérdida.

La calificación que otorgan las compañías calificadoras en general es:

1. La opinión de incurrir en incumplimiento.
2. Una medida comparable, globalmente, de riesgo de incumplimiento entre un amplio rango de posibles acreditados.
3. Un pronóstico del grado relativo de protección que un inversionista tiene si un acreditado enfrenta una adversidad económica.
4. La evaluación de la capacidad y compromiso jurídico de un acreditado para efectuar los pagos de intereses y amortizaciones del principal, en el plazo determinado en las condiciones de emisión del título de renta fija específico.
5. Una medida de probabilidad de que un acreditado incurra en incumplimiento con respecto a sus obligaciones en el contrato de crédito durante la duración del mismo. La duración depende del instrumento, puede ser de unos cuantos días o, hasta de 30 años o más. Asimismo, las calificaciones a largo plazo incorporan una evaluación de cuál es la pérdida monetaria en caso de incumplimiento.

Es importante mencionar que esta calificación no representa:

1. Una predicción de la estabilidad o volatilidad del precio de un título de deuda.

2. Un juicio individual de la calidad de la administración.
3. Elogios o críticas a una determinada estrategia de administración de pasivos.
4. Una recomendación para comprar o vender.
5. Sólo el resultado de razones financieras.
6. Una garantía en contra de pérdidas.
7. Estas calificaciones se enfocan en el riesgo de pérdida crediticia debido a un pago que no se realiza o que se retrasa. Las calificaciones no están destinadas a medir otros tipos de riesgos que pueden estar asociados con las inversiones de renta fija, tales como el riesgo de pérdida en el valor de mercado de un instrumento debido a fluctuaciones en las tasas de cambio de divisas, en las tasas de interés, o debido al reintegro del capital de un préstamo antes del vencimiento.
8. Además, a diferencia de las calificaciones de acciones, las calificaciones de deuda no están destinadas a medir el potencial de apreciación de un valor.

La emisión de las calificaciones genera beneficios para los inversionistas y los acreditados.

Los beneficios para un inversionista son:

1. Reduce la incertidumbre
 - a) Fomenta el crecimiento del mercado de capital.
 - b) Aumenta la eficiencia del mercado y la liquidez.
2. Utiliza las calificaciones como un elemento clave para establecer las primas de riesgo en función del riesgo crediticio para los títulos de compra. Es decir, utiliza las calificaciones para medir el rendimiento adicional que debe exigir el inversionista para poder ser compensado por la posible pérdida crediticia estimada sobre el título que se está comprando.
3. Amplía los horizontes de inversión, pues es una evaluación independiente y objetiva del futuro riesgo crediticio a largo plazo, de acuerdo con una norma comparable a escala mundial.

- a) Es consistente de manera global y con estándares creíbles.
- b) Permite la diversificación del portafolio.

Los beneficios para el acreditado son:

1. Aumenta la flexibilidad del financiamiento.
2. Incrementa el acceso a los mercados de capital.
3. Las calificaciones afectan a la liquidez (estabilizando el acceso al mercado).

3.1.2. Importancia de las Calificaciones de Riesgo de Crédito

Las calificaciones de riesgo de crédito son importantes para los mercados de capitales por diversos motivos:

1. Contribuyen a una asignación más eficiente de los ahorros y la inversión, lo que a su vez promueve un más rápido crecimiento económico. Al contar con una medición confiable del riesgo de incumplimiento, los inversionistas pueden balancear más fácilmente su tolerancia al riesgo y sus objetivos de inversión.

En América Latina y el Caribe, regiones con apremiantes necesidades de inversión, y donde los ahorros son, en general, un recurso escaso, las calificaciones crediticias pueden contribuir a asegurar que dichos ahorros de la región sean utilizados más eficientemente.

2. Las calificaciones crediticias reducen el riesgo de información y de este modo las primas de riesgo exigidas por los inversionistas, ya que, los inversionistas que operan en un entorno donde existe escasa información financiera sobre estrategia corporativa, resultados financieros o posición de negocios, exigirán mayores tasas de rendimiento sobre sus inversiones.

Las calificaciones crediticias ayudan a los inversores a realizar una mejor evaluación y comparación del riesgo crediticio a través de un amplio

espectro de compañías, sectores industriales y países, contribuyendo a mantener estándares de información y análisis financieros orientados a disminuir el costo de capital, lo que a su vez estimula la inversión y promueve el crecimiento económico.

3. Las calificaciones crediticias fomentan mercados financieros más líquidos. En todas partes del mundo, los inversionistas desean alternativas para sus ahorros. Si los mercados de capitales locales son transparentes y brindan buenos retornos ponderados por riesgo, los inversionistas cuentan con mayores incentivos para colocar su dinero localmente.

Por el contrario, falta de liquidez o transparencia alientan a los inversionistas a colocar su dinero *offshore* o en activos reales. Mercados financieros más profundos también fomentan políticas monetarias prudentes, y permiten a los gobiernos emitir deuda dirigida a un mayor número de inversionistas.

Un beneficio importante de dichas iniciativas es que contribuyen a que los bonos del gobierno emitidos localmente actúen como un *benchmark* para otros deudores del sector público o privado.

4. Las calificaciones crediticias señalan la intención del deudor de ser abierto y transparente con sus acreedores ya que, obtener y mantener una calificación implica someterse a un proceso de provisión de información exhaustiva. Los inversionistas cuentan con el compromiso del equipo gerencial de realizar este proceso, así como también con el de la calificadora de riesgo de mantener su opinión actualizada.

3.2. Matrices de Transición

El proceso por medio del cual un acreditado migra de una calificación a otra, puede ser modelado matemáticamente como una cadena de *Markov* finita. En ésta se asume que, dado un intervalo de tiempo, las calificaciones de crédito cambian sucesivamente de un estado a otro, con una cierta probabilidad. Las probabilidades de migración de crédito forman una matriz de transición.

3.2.1. Cadenas de Markov y Matrices de Transición

Sea (Ω, σ, P) un espacio de probabilidad y $E \subset R$ un subconjunto finito o numerable. Una sucesión de variables aleatorias

$$\{X_n : \Omega \rightarrow E; n = 0, 1, 2, \dots\}$$

se llama cadena de *Markov* con espacio de estados E si satisface la condición de *Markov*, esto es, si para todo $n \geq 1$ y toda sucesión $x_0, x_1, \dots, x_{n-2}, x, y \in E$ se cumple

$$P(X_n = y | X_{n-1} = x, \dots, X_0 = x_0) = P(X_n = y | X_{n-1} = x) \quad (3.1)$$

donde la distribución de X_0 se llama distribución inicial. Denotaremos por π al vector $\pi = (P(X_0 = x), x \in E)$.

Las cadenas de *Markov* cumplen con:

1. El lado derecho la ecuación (3.1) determina que la probabilidad de estar en el estado y al tiempo n dado que en los tiempos anteriores la cadena siguió la trayectoria $\{x_0, \dots, x_{n-2}, x\}$, sólo depende del estado inmediato anterior, es decir, del estado al tiempo $n - 1$.
2. Si $P(X_n = y | X_{n-1} = x)$ no depende de n , es decir, es independiente del tiempo, entonces la cadena es homogénea (con respecto al tiempo). Para una cadena de *Markov* homogénea denotaremos a $P(X_n = y | X_{n-1} = x)$ por $P_{x,y}$. La familia $\{P_{x,y} : x, y \in E\}$ se le llama las probabilidades de transición de la cadena. En ésta se describe la evolución de la cadena en el tiempo.
3. Para $m \geq 1$ se denota por $P_{x,y}^{(m)}$ a $P(X_{n+m} = y | X_n = x)$, y representa la probabilidad de ir en m unidades de tiempo de x a y .
4. Para $x, y \in E$ se define a $P_{x,y}^{(0)}$ como $\delta_{x,y}$ donde

$$\delta_{x,y} = \begin{cases} 1 & \text{si } x = y \\ 0 & \text{si } x \neq y \end{cases}$$

5. La matriz $P = (P_{x,y})_{x,y \in E}$ es llamada matriz de transición.

Ahora vamos a ver que la matriz de transición es una matriz estocástica, es decir, una matriz con entradas no negativas y tal que la suma de las entradas por renglón es uno.

Como $P_{x,y}$ es una probabilidad fija para cada pareja $x, y \in E$, entonces $P_{x,y} \geq 0$, y por tanto la matriz de transición tiene todas sus entradas no negativas.

Para ver que la suma de sus entradas por renglón es uno, tomando las propiedades de la imagen inversa de una función se tiene que:

$$\Omega = \{\omega \in \Omega | X_n(\omega) \in E\} = \cup_{x \in E} \{\omega \in \Omega | X_n(\omega) = x\}$$

siempre que $X_n : \Omega \rightarrow E$ sea una función y se trata de una unión de conjuntos ajenos dos a dos. Además,

$$\Omega = \cup_{x_1, \dots, x_{n-1} \in E} \{\omega \in \Omega | X_{n-1}(\omega) = x_{n-1}, \dots, X_0 = x_0\}$$

siempre que $X_i : \Omega \rightarrow E$ con $i = 1, 2, \dots, n-1$ sean funciones y nuevamente son conjuntos ajenos dos a dos.

Lema 3.2.1 Para todo $x \in E$ se tiene que $\sum_{y \in E} P_{x,y} = 1$.

Demostración:

Sea $x \in E$,

$$1 = P_x(X_{n+1} \in E) = P(X_{n+1} \in E | X_n = x) = \sum_{y \in E} P_{x,y} \blacksquare$$

En la modelación del riesgo de crédito, una matriz de transición da la información acerca de la probabilidad de transición de una calificación a otra. Además, las compañías calificadoras asumen independencia del tiempo a las

probabilidades de las matrices de transición. Por estas razones, en la práctica se tiene que para simular eventos de incumplimiento se tienen cadenas de *Markov* homogéneas.

Para matrices de transición en riesgo de crédito se tiene que E es finito, es decir, $E = 1, 2, \dots, n$

Proposición 3.2.1 Ecuación de Chapman-Kolmogorov.

Para una cadena de *Markov* homogénea $(X_n)_{n=0,1,\dots}$ con espacio de estados E y para todo $n, m \in \mathbb{N}$ y toda pareja $x, y \in E$ se cumple

$$P(X_{n+m} = y | X_0 = x) = \sum_{z \in E} P_{x,z}^{(n)} P_{z,y}^{(m)} = P_{x,y}^{(n+m)} \quad (3.2)$$

Demostración:

Se hará por inducción. Se supone que para todo $k \in \mathbb{N}$ con $k < (n + m)$ se cumple (3.2). Ahora veremos que se cumple para $(n + m)$:

$$\begin{aligned} P(X_{n+m} = y | X_0 = x) &= \frac{P(X_{n+m} = y, X_0 = x)}{P(X_0 = x)} \\ &= \frac{P(X_{n+m} = y, \cup_{z \in E} X_n = z, X_0 = x)}{P(X_0 = x)} \\ &= \sum_{z \in E} \frac{P(X_{n+m} = y, X_n = z, X_0 = x)}{P(X_0 = x)} \\ &= \sum_{z \in E} \frac{P(X_{n+m} = y | X_n = z, X_0 = x) * P(X_n = z, X_0 = x)}{P(X_0 = x)} \\ &= \sum_{z \in E} P(X_{n+m} = y | X_n = z) * P(X_n = z | X_0 = x) \\ &= \sum_{z \in E} P_{x,z}^{(n)} P_{z,y}^{(m)} = P_{x,y}^{(n+m)} \blacksquare \end{aligned}$$

Como E es finito, $P_{x,y}^{(k)}$ es la entrada x, y de la potencia k -ésima de la matriz de transición P .

Entonces tenemos que:

$$P(t) = P^t \quad (3.3)$$

De la ecuación (3.3) tenemos que la distribución de probabilidad de la cadena de Markov al tiempo t satisface:

$$\pi(t) = \pi(0)P^t = \pi P^t$$

donde

$$\pi(t) = (P(X_t = x), x \in E).$$

Clases de Comunicación y Clasificación de Estados

En una cadena de Markov $(X_n)_{n \in \mathbb{N}}$ con espacio de estados E y matriz de transición $(P_{x,y})_{x,y \in E}$ se tiene que:

1. $x, y \in E$ se comunica entre sí, y se denota $(x \longleftrightarrow y)$, si se cumplen:
 - a) x se comunica con y , $(x \longrightarrow y)$, es decir, existe $n \geq 0$ tal que $P_{x,y}^{(n)} > 0$.
 - b) y se comunica con x , $(y \longrightarrow x)$, es decir, existe $m \geq 0$ tal que $P_{x,y}^{(m)} > 0$.
2. Decimos que $C \subset E$ es una clase comunicante si:

$$C = C(x) = \{y \in E : x \longleftrightarrow y\}.$$

3. Si la clase comunicante es única, es decir si $C(x) = E$ para algún $x \in E$, entonces a la cadena se le llama irreducible o ergódica.

4. Un estado $x \in E$ es estado recurrente si:

$$P(X_n = x \text{ para alguna } n \in \mathbb{N} | X_0 = x) \equiv 1.$$

5. Un estado $x \in E$ es estado transitorio si:

$$P(X_n = x \text{ para alguna } n \in \mathbb{N} | X_0 = x) < 1.$$

6. Un estado $x \in E$ es estado absorbente si:

$$P_{x,x} \equiv 1.$$

Es decir, una vez que se llega a este estado, ya nunca se sale de él.

3.2.2. La Matriz de Transición de Calificaciones Crediticias

Las compañías calificadoras toman a las clasificaciones como el espacio de estados y generan la correspondiente matriz de transición.

Una matriz de transición de este tipo es de la forma:

	AAA	AA	A	BBB	BB	B	C	D
AAA	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,5}$	$x_{1,6}$	$x_{1,7}$	$x_{1,8}$
AA	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	$x_{2,5}$	$x_{2,6}$	$x_{2,7}$	$x_{2,8}$
A	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$x_{3,4}$	$x_{3,5}$	$x_{3,6}$	$x_{3,7}$	$x_{3,8}$
BBB	$x_{4,1}$	$x_{4,2}$	$x_{4,3}$	$x_{4,4}$	$x_{4,5}$	$x_{4,6}$	$x_{4,7}$	$x_{4,8}$
BB	$x_{5,1}$	$x_{5,2}$	$x_{5,3}$	$x_{5,4}$	$x_{5,5}$	$x_{5,6}$	$x_{5,7}$	$x_{5,8}$
B	$x_{6,1}$	$x_{6,2}$	$x_{6,3}$	$x_{6,4}$	$x_{6,5}$	$x_{6,6}$	$x_{6,7}$	$x_{6,8}$
C	$x_{7,1}$	$x_{7,2}$	$x_{7,3}$	$x_{7,4}$	$x_{7,5}$	$x_{7,6}$	$x_{7,7}$	$x_{7,8}$
D	$x_{8,1}$	$x_{8,2}$	$x_{8,3}$	$x_{8,4}$	$x_{8,5}$	$x_{8,6}$	$x_{8,7}$	$x_{8,8}$

donde cada $x_{i,j}$ representa la probabilidad de pasar de la calificación i a la j o se mantenga en la misma en un periodo de tiempo establecido.

Y entonces el espacio de estados es $E = \{AAA, AA, A, BBB, BB, B, C, D\}$.
Con $\sum_{j=1}^8 x_{i,j} = 1$ para toda $i = 1, 2, \dots, 8$ y $x_{i,j} \geq 0$ para toda $i, j = 1, 2, \dots, 8$.

Los modelos de riesgo de crédito usualmente suponen que, con el paso del tiempo, los acreditados, emigran de una calificación crediticia a otra.

En un determinado periodo de tiempo $T = \Delta t = t_1 - t_0, 2\Delta t, \dots$ se da seguimiento al movimiento de la calidad crediticia de C_j a cualquier otra C_k por acreditado donde cada C_i es una calificación.

Si denotamos $C_{i,n}$ a la calificación del acreditado i al tiempo n con $C_{i,n} \in E$ y $n \in N$.

Entonces se tienen las siguientes hipótesis:

1. Todos los acreditados tienen características homogéneas sin considerar la misma calificación.

Es decir:

Si $C_{i,n} = C_{j,n}$ entonces

$$\implies P[C_{i,n} = C_k] = P[C_{j,n} = C_k] \quad \forall i, j, k$$

2. La probabilidad de transición de cada uno de los acreditados depende sólo de la calificación inmediata anterior.

Entonces:

$$P[C_{i,n+1} = C_{k(n+1)} | C_{j,n} = C_{k(n)}] = P[C_{i,n+1} = C_{k(n+1)} | C_{j,n} = C_{k(n)}, \dots, C_{j,0} = C_{k(0)}]$$

3. Las probabilidades de transición son estacionarias, es decir, no dependen del tiempo.

Entonces:

$$P[C_{i,n+1} = C_{k(n+1)} \mid C_{j,n} = C_{k(n)}] = P[C_{i,n} = C_{k(n)} \mid C_{j,n-1} = C_{k(n-1)}]$$

Puesto que en la matriz de transición se reflejan las probabilidades en la calidad de migración, ésta es una herramienta que permite modelar el cambio en la calidad de crédito. Además, con base en estas matrices de transición es posible generar los diferentes escenarios en la calidad de un portafolio de un periodo de tiempo a otro.

Debido a que las matrices de transición pueden ser obtenidas mediante diferentes métodos, resulta bastante importante tener un criterio para discriminar a una matriz de otra. En la siguiente sección se propone una metodología para la discriminación de dichas matrices.

3.3. Métodos de Estimación de Matrices de Transición

En esta sección se presentan los métodos que serán utilizados para estimar las matrices de migración. La estimación está basada en las calificaciones históricas de crédito.

3.3.1. Modelo Regular de Migración de Crédito

Las cadenas de *Markov* utilizadas en los modelos de riesgo de crédito tienen un estado especial que representa el evento de incumplimiento. Este estado es absorbente y se le asigna el estado n .

Entonces tenemos que:

$$P_{n,n}(t) = 1 \text{ y } P_{n,j}(t) = 0 \text{ para todo } j \neq n \text{ y } t = 1, 2, \dots$$

Definiendo al Modelo Regular de Migración de Crédito (MRMC) como una cadena de *Markov* que satisface las siguientes condiciones:

1. Existe uno y sólo un estado absorbente.

2. Existe $t_p \geq 1$ tal que $P_{i,n}(t_p) > 0$ para todo $i = 1, 2, \dots, n$.
3. El determinante de la matriz P es diferente de cero, y sus eigenvalores son distintos.

Además para el (MRMC), la matriz de transición $P(T)$ satisface (Snell 1988)

$$P(t) \rightarrow D \quad (3.4)$$

cuando $t \rightarrow \infty$

donde:

$$D = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

La ecuación (3.4) implica que el incumplimiento eventualmente ocurre independientemente de la calificación inicial de crédito. Todos los estados se comunican con el estado de incumplimiento, es decir, $(x \rightarrow n)$, para todo $x = 1, 2, \dots, (n-1)$.

El problema

Dada una cadena de Markov $(X_t)_{t \in \mathbb{N}}$ finita en tiempo discreto el problema puede formularse como:

La posibilidad de construir un proceso de Markov $\{\tilde{X}_\tau, \tau \geq 0\}$ en tiempo continuo tal que la distribución de probabilidad de \tilde{X}_τ al tiempo $\tau = 1, 2, \dots$ es igual a la distribución de (X_t) .

Este problema fue estudiado por Kingman (1962), Carrette (1995), Iwanik y Shiflett (1986), y Frydman (1993)[1]. Aquí la principal propuesta desarrollada es la basada en la noción de un generador de un proceso de Markov.

Un generador es una matriz G cuyos elementos satisfacen:

$$\sum_{j=1}^n g_{i,j} = 0 \quad (3.5)$$

para $i = 1, 2, \dots, n$. Y $g_{i,j} \geq 0$ para $i, j = 1, 2, \dots, n$ y $i \neq j$.

Los generadores forman un semigrupo en el espacio de matrices, lo cual significa que la suma de dos generadores es también un generador.

Por ejemplo si P es la matriz de transición entonces $P - I$ es un generador, donde I es la matriz identidad de iguales dimensiones que la matriz P . Además, si G es un generador entonces $\exp(G)$ es una matriz de transición Snell (1988).

Sea $\hat{P}(\tau)$ la matriz de transición del proceso de Markov \hat{X}_τ .

La matriz $\hat{G} = (\hat{g}_{i,j})_{i,j \in E}$ es llamada la matriz generadora de \hat{X}_τ si $\hat{P}(\tau)$ satisface:

$$\frac{d}{d\tau} \hat{P} = \hat{P} \hat{G}.$$

Entonces tenemos que:

$$\hat{P}(\tau) = \exp(\tau \hat{G})$$

con $\tau \geq 0$.

Si la distribución del proceso de Markov \hat{X}_τ es idéntica a la de la cadena de Markov X_t al tiempo $\tau = 1, 2, \dots$ entonces, la matriz de transición P debe de satisfacer:

$$P = \exp(\hat{G}). \quad (3.6)$$

Si la matriz \hat{G} existe y es única, entonces se puede definir la extensión al tiempo continuo, $\hat{P}(\tau)$ de la matriz de transición P como:

$$\hat{P}(\tau) = \exp(\tau \hat{G}). \quad (3.7)$$

Entonces como τ es continua, de la ecuación (3.7) podríamos obtener matrices de transición para cualquier periodo arbitrario de tiempo.

Calculando $\hat{G} = \ln(P)$ de la ecuación (3.6), no se garantiza que la matriz $\ln(P)$ satisfaga la ecuación (3.5). En general, saber cuando una matriz de transición P puede ser representada en forma de $P = \exp(G)$, donde G es

un generador, es técnicamente muy difícil.

La propuesta que se plantean *Alexander Kreinin y Marina Sidelnikova* está basada en la descomposición de *Schur* y el método *Parleet* (*Moler y Van Loan 1978*)[1].

La No-Unicidad del Problema

Aquí se muestra el problema de la no-unicidad de la matriz resultante de transición cuando se calculan probabilidades de transición de crédito para horizontes menores a un año.

Un ejemplo ilustrativo muestra este hecho:

Considerando la familia de matrices

$$P(0,5) = \begin{pmatrix} 0 & p & \dots & 0 & 0 & 1-p \\ x & 0 & \dots & 0 & 0 & 1-x \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & p & 1-p \\ 0 & 0 & \dots & 0 & x & 1-x \\ 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}.$$

Imponiendo la condición que $px = a$, donde $0 < p < 1$ y $0 < x < 1$. Entonces, la matriz de transición anual, $P = P(0,5) * P(0,5)$ es:

$$P(0,5) = \begin{pmatrix} a & 0 & 0 & \dots & 0 & 1-a \\ 0 & a & 0 & \dots & 0 & 1-a \\ 0 & 0 & a & \dots & 0 & 1-a \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a & 1-a \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Pero, existe otra matriz de transición $P^*(0,5)$ igual a la raíz de P :

$$P^*(0,5) = \begin{pmatrix} \sqrt{a} & 0 & 0 & \dots & 0 & 1-\sqrt{a} \\ 0 & \sqrt{a} & 0 & \dots & 0 & 1-\sqrt{a} \\ 0 & 0 & \sqrt{a} & \dots & 0 & 1-\sqrt{a} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \sqrt{a} & 1-\sqrt{a} \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

Note que $P^*(0,5)$ es la solución que se obtiene usando el generador de la sección anterior.

La pregunta natural ahora es:

¿Qué condiciones se deben cumplir para que la matriz de transición resultante sea única? Una respuesta parcial puede ser encontrada en los siguientes proposiciones (Israel (2001)).

Proposición 3.3.1 *Supongase que todos los elementos de la diagonal de la matriz de transición P dominan, es decir, $p_{i,i} \geq \frac{1}{2}$. Entonces los eigenvalores de P pertenecen al disco $\|\lambda - 1\| \leq 1$ en el plano complejo, y la serie de expansión de Taylor para el logaritmo de la matriz de transición converge geoméricamente rápidamente a una matriz real Q , que satisface $\exp(Q) = P$. Además bajo estas condiciones, la matriz de transición puede tener, a lo más, un generador.*

Demostración[1].

Cabe mencionar que la proposición no dice que el hecho de que la diagonal domine garantiza la existencia del generador.

Proposición 3.3.1 *Supongase que la matriz de transición P , satisface una de las siguientes condiciones:*

1. $\det(P) < 0$.
2. $\det(P) > \prod_{i=1}^n p_{i,i}$.
3. *Existen estados i y j tales que i se comunica con j , pero $p_{i,j} = 0$.*

Entonces no existe un generador exacto de P .

Demostración[1].

La tercera condición se satisface en la mayoría de las matrices de transición de riesgo de crédito. Lo cual es una explicación del porque esta condición debe cumplirse para un (MRMC).

Dado que en las matrices de transición de riesgo de crédito todos los estados se comunican con el estado de incumplimiento y $p_{i,n} = 0$ para todo $i = 1, 2, \dots, (n - 1)$, es decir se cumple la tercera condición del teorema 2; entonces podemos concluir que en la mayoría de los casos de este tipo de matrices el generador exacto no existe.

Tres Alternativas

Asumiendo homogeneidad de tiempo, el principal problema es encontrar una matriz de transición X tal que:

$$X^t = P$$

con $t > 1$ y donde P es la matriz de transición anual y t es el número de periodos por un año.

Sea $MT(n)$ el conjunto de todas las matrices de transición de dimensión $n \times n$ que satisfacen la condición de matriz estocástica.

Calculando $X = P^{\frac{1}{t}} = \exp(\frac{1}{t} \ln(P))$ podría resultar una matriz con entradas negativa y entonces X no pertenecería a $MT(n)$. Si existe un generador \hat{G} tal que $\exp(\hat{G}) = P$, entonces $\hat{X} = \exp(\frac{1}{t} \hat{G})$ es un elemento de $MT(n)$.

El problema es que en la mayoría de los casos, las matrices de transición anuales de riesgo de crédito no tienen generador.

La regularización del problema puede ser descrita por medio de la siguiente metodología:

Primero encontrar una matriz de transición X tal que, elevada a la potencia t sea la más cercana a la matriz anual de transición P .

Matemáticamente el problema puede formularse como:

Problema (MAM) Mejor Aproximación a la Matriz anual de transición.

Encontrar $\hat{X} \in MT(n)$ tal que:

$$\| \hat{X}^t - P \| = \min_{X \in MT(n)} \| X^t - P \|$$

donde $\| * \|$ es la distancia en el espacio de las matrices de $n \times n$.

Problema de Optimización de la Matriz (OM)

Encontrar una $\hat{X} \in MT(n)$ tal que:

$$\| \hat{X} - P^{\frac{1}{t}} \| = \min_{X \in MT(n)} \| X - P^{\frac{1}{t}} \|$$

Aquí el objetivo es encontrar la matriz de transición lo más aproximada posible a la matriz fraccional anual de transición $P^{\frac{1}{t}} = \exp(\frac{1}{t} \ln(P))$.

Para resolver *OM*, se utiliza el hecho de que el conjunto de las matrices, $MT(n)$, puede ser representado como un producto cartesiano n -dimensional de n vectores. Además como cada renglón de la matriz de transición satisface el Lema 4.1 y pertenecen a los vectores n -dimensionales, $Vec(n)$, definidos como:

$$Vec(n) = \{(x_1, \dots, x_n) \in R^n, \sum_{i=1}^n x_i = 1, x_i \geq 0\} \quad (3.8)$$

Además $Vec(n)$ está contenido en el hiperplano $H(n)$

$$H(n) = \{(x_1, \dots, x_n) \in R^n, \sum_{i=1}^n x_i = 1\}$$

Utilizando la norma *Euclidiana* para medir la distancia entre cualesquiera dos puntos $x, y \in R^n$, es decir:

$$dist(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2}$$

Entonces el problema *OM* puede resolverse renglón por renglón, es decir puede ser reducido a n independientes instancias del siguiente problema de

minimización de distancias:

Dado un punto $p \in R^n$, $p = (p_1, \dots, p_n)$, encontrar $x \in Vec(n)$ tal que:

$$dist(p, x) = \min_{y \in Vec(n)} dist(p, y)$$

El problema puede ser resuelto siguiendo el siguiente algoritmo, el cual fué sugerido por *Merkoulovitch*(2000)[1].

1. Encontrar la proyección b del punto p en el hiperplano $H(n)$, con $b_i = p_i - \lambda$, donde $\lambda = \frac{1}{n}(\sum_{i=1}^n p_i - 1)$.
2. Si todas las coordenadas de b son no negativas entonces paramos y b es la solución al problema.
3. Sea $\hat{p} = \pi(b)$, donde π es una permutación que ordena las coordenadas de b en forma descendiente.
4. Calcular $c_k = \sum_{i=1}^k \hat{p}_i - k\hat{p}_k$ con $k = 1, 2, \dots, n$. Las sumas de c_k satisfacen $0 \leq c_1 \leq c_2 \leq \dots \leq c_n$.
5. Encontrar $k^* = \max\{k : k \geq 1, c_k \leq 1\}$
6. Construir el vector $\hat{x} \in vec(n)$ de la siguiente manera.

$$\hat{x}_j = \begin{cases} 0 & \text{si } j > k^* \\ \hat{p}_j + \frac{1}{k^*}(1 - \sum_{i=1}^{k^*} \hat{p}_i) & \text{si } j \leq k^* \end{cases}$$

7. Aplicar la permutación inversa π^{-1} a \hat{x} ; y entonces $\pi^{-1}(\hat{x})$ es la solución al problema.

Este algoritmo cumple con lo siguiente:

Proposición 3.3.1 Sea $p = (p_1, \dots, p_n)$ el punto inicial y sea $x^* = (x_1^*, \dots, x_n^*)$ la solución óptima. Si $p_1 \geq \dots \geq p_n$ entonces $x_1^* \geq \dots \geq x_n^*$.

Demostración [1].

Esta proposición establece que los elementos de la solución óptima están ordenados con la misma secuencia que los del punto inicial.

Proposición 3.3.1 Si b es la proyección de p en $H(n)$ y $b_k < 0$ para alguna k , entonces $x_j^* = 0$ para $j=k, \dots, n$.

Demostración [1]

En esta proposición se establece que si después de la proyección en el hiperplano, algunas de las coordenadas son negativas, entonces, en la solución óptima éstas coordenadas son iguales a cero. Lo cual permite reducir el problema original a uno de optimización discreta como:

Con λ obtenida como en el paso 1 del algoritmo, defínase la función $f(l) = l * \lambda^2 + \sum_{i=l+1}^n a_i^2$ para $l = 1, 2, \dots, n$. Notar que $\sum_{i=k}^m a_i^2 \equiv 0$ para $k > m$.

La solución del problema de minimización de distancia puede ser obtenida resolviendo:

$\min f(l)$
sujeto a

$$l * \lambda + \sum_{i=1}^l a_i = 1$$

$$\lambda + a_l \geq 0$$

$$l = 1, \dots, n.$$

La solución l^* a este problema determina el número óptimo de coordenadas de k^* iguales a cero en el paso 5 del algoritmo.

Proposición 3.3.1 *La función objetivo $f(l)$ es monótona, es decir, $f(l) \geq f(l+1)$.*

Demostración:

La proposición se sigue de la identidad

$$f(m) = f(m+1) + \frac{1}{m * (m+1)} * (1 - T_m)^2$$

donde

$$T_m = \sum_{i=1}^m a_i - m * a_{m+1} \blacksquare$$

De esta proposición se sigue que la solución óptima del problema es:

$$l^* = \max\{l : 1 \leq l \leq n, l * a_l \geq \sum_{i=1}^l a_i - 1\}$$

Esto produce un l^* igual a k^* , calculado en el paso 5 del algoritmo.

Además, el problema de minimización de distancia puede ser resuelto de manera iterativa reemplazando el paso 3 del algoritmo por:

Paso 3* . Fijar cualquier elemento negativo de b igual a cero, hacer $p = b$ y regresar al paso 1.

Este algoritmo iterativo termina después de m pasos donde m no excede el tamaño del vector p .

Problema de la Optimización de la Matriz Generadora (OMG)

Definiendo el conjunto de las matrices generadoras $G(n)$, como el conjunto de todas las matrices de dimensión $n \times n$ que satisfacen la ecuación 3.5.

El problema consiste en encontrar $\hat{G} \in G(n)$ tal que:

$$\| \hat{G} - \ln(P) \| = \min_{X \in G(n)} \| X - \ln(P) \|$$

El problema *OMG* es diferente que el problema *OM* desde el punto de vista geométrico. Mientras el espacio de matrices de transición $MT(n)$ es un producto cartesiano compuesto por n vectores, el espacio de los generadores, $G(n)$, es un producto cartesiano de conos n -dimensionales. Cada renglón de un generador tiene la propiedad que sus elementos suman cero y los elementos fuera de la diagonal son no negativos. Permutando los elementos de un renglón, siempre se pueden representar como un punto en el cono estandar, $C(n)$ definido por:

$$C(n) = \{(x_1, \dots, x_n) \in R^n, \sum_{i=0}^n x_i = 0, x_i \geq 0, i \geq 2\}$$

Además $C(n)$ está contenido en el hiperplano $\hat{H}(n)$:

$$\hat{H}(n) = \{(x_1, \dots, x_n) \in R^n, \sum_{i=0}^n x_i = 0\}$$

De manera similar al problema *OM*, el problema *OMG* puede resolverse renglón por renglón basándose una vez más en proyectar un punto $p \in R^n$, es decir un vector de la matriz $\ln(P)$.

Entonces, el problema (*OMG*) puede también resolverse con n instancias independientes con el siguiente problema de minimización de distancias:

Dado un punto $p \in R^n$, $p = (p_1, \dots, p_n)$, encontrar $g^* \in C(n)$ tal que:

$$\text{dis}(p, g^*) = \min_{g \in C(n)} \text{dis}(p, g)$$

La solución óptima a este problema puede ser obtenida por el siguiente algoritmo:

1. Sea b la proyección de p en $\hat{H}(n)$. Con $b_i = p_i - \lambda$, donde $\lambda = \frac{1}{n}(\sum_{i=1}^n p_i)$.

2. Sea $\hat{p} = \pi(b)$, donde π es una permutación que ordena las coordenadas de b en forma descendente.
3. Encontrar l^* , el menor entero $2 \leq l \leq n-1$ que satisfaga:

$$(n-l+1)\hat{p}_{l+1} \geq \hat{p}_l + \sum_{i=0}^{n-(l+1)} \hat{p}_{n-i}$$

4. Sea $\xi = \{i : 2 \leq i \leq l^*\}$. Construir el vector $\hat{g} \in C(n)$ como:

$$\hat{g}_i = \begin{cases} 0 & \text{si } i \in \xi \\ \hat{p}_j + \frac{1}{(n-l^*+1)} (\sum_{j \notin \xi} \hat{p}_j) & \text{en otro caso} \end{cases}$$

5. Aplicar la permutación inversa π^{-1} a \hat{g} ; entonces, $\pi^{-1}(\hat{g})$ es la solución al problema de minimización de distancias.

Otros Métodos para los Generadores

Los siguientes métodos fueron desarrollados por *Stromquist*(1996) y *Araten y Angbazo* (1997). Estos métodos ajustan la matriz $G = \ln(P)$, para construir un generador \hat{G} .

Para satisfacer la ecuación 3.7, todos los elementos fuera de la diagonal que sean negativos en G se igualan a cero y se ajustan los renglones para que sumen cero.

Existen dos métodos para realizar este ajuste, uno llamado Ajuste de Diagonal AD , que ajusta sólo elementos elementos de la diagonal y otro, Ajuste Completo AC , que ajusta a todos los elementos distintos de cero.

Para calcular \hat{G} se siguen los siguientes pasos:

1. Sea

$$\hat{g}_{i,j} = \begin{cases} 0 & \text{si } i \neq j \text{ y } g_{i,j} < 0 \\ q_{i,j} & \text{en otro caso} \end{cases}$$

para $i, j = 1, \dots, n$.

2. (Ajuste de Diagonal). Se ajustan los elementos de la diagonal con respecto a los que no están en la diagonal de manera que la suma de los elementos por renglón sea cero. El ajuste está definido por:

$$\hat{g}_{i,i} = - \sum_{j=1, j \neq i}^n \hat{g}_{i,j}$$

para $i, j = 1, 2, \dots, n$ o bien.

3. (Ajuste Completo). Ajustando todos los elementos distintos de cero de acuerdo a sus magnitudes relativas:

$$\hat{g}_{i,i} = \hat{g}_{i,i} - |\hat{g}_{i,i}| \frac{\sum_{i=1}^n \hat{g}_{i,i}}{\sum_{i=1}^n |\hat{g}_{i,i}|}$$

para $i, j = 1, 2, \dots, n$.

3.3.2. Método *Cohort*

Sea $p_{i,j}(\Delta)$ la probabilidad de pasar de la calificación i a la j en el tiempo Δt . Por otro lado, sea n_i el número de acreditados con calificación i al inicio del periodo Δt , $n_{i,j}$ el número de acreditados que pasaron de la calificación i a la j para el final del periodo de tiempo Δt . Entonces una estimación de la probabilidad de transición es:

$$p_{i,j}(\Delta) = \frac{n_{i,j}}{n_i} \quad (3.9)$$

Cabe mencionar que con esta probabilidad estimada cualquier cambio ocurrido en el transcurso del periodo es ignorado.

3.3.3. Método de Duración o Tasa de Riesgo

En las calificaciones históricas de crédito existen dos aspectos muy importantes: el análisis sensorial, es decir, saber que ocurre con los acreditados entre cada periodo que reciben una calificación y el hecho de que los posibles acreditados sólo son tomados en cuenta si tienen un cierto tiempo de existencia o

tienen una calificación. Estos dos aspectos son ignorados en el método *cohort*.

La construcción formal de este método es una cadena de *Markov* homogénea con N estados, donde el estado 1 representa la mejor calificación y van en orden descendente hasta el estado N que es el estado de incumplimiento.

Con la homogeneidad de tiempo, las probabilidades de transición pueden describirse por medio de una matriz generadora G de $N \times N$. La matriz de probabilidad de transición $P(t)$, siguiendo los resultados de *Lando y Skodeberg*(2002), está dada por:

El estimador de máxima verosimilitud de $g_{i,j}$ está dado por:

$$\hat{g}_{i,j} = \frac{n_{i,j}(T)}{\int_0^T Y_i(s) ds} \quad (3.10)$$

donde $Y_i(s)$ es el número de firmas con calificación i al tiempo s , y $n_{i,j}(T)$ es el número total de transiciones de i a j en el periodo con $i \neq j$. Entonces, el denominador efectivamente es el número de firmas enviadas en el estado i en el periodo de tiempo.

3.3.4. Máxima Entropía

Supongamos que tenemos K acreditados para un periodo t y un espacio de estados(calificaciones) finito $j = 1, 2, \dots, n$.

Sea $x_{i,j,t} = 1$ una variable aleatoria, si el acreditado i -ésimo con $i = 1, 2, \dots, K$, está en la calificación j -ésimo al tiempo t y $x_{ij,t} = 0$ para todos los $n - 1$ calificaciones.

Definiendo X_t como el vector $n - dimensional$ de proporciones del j -ésimo en el periodo t . Entonces, los elementos de X_t representan repentan la proporción de acreditados en cada calificación, al tiempo t . Análogamente, X_{t+1} es el $n - dimensional$ vector de proporciones de acreditados en la $j - esimo$ calificación al tiempo $t + 1$.

El objetivo es estimar una matriz de probabilidades de transición de calificación crediticia P de $n \times n$ utilizando la información histórica.

Entonces, tenemos una relación lineal entre X_t y X_{t+1} y la matriz de probabilidades de transición P dada por:

$$x_{t+1,j} = \sum_{k=1}^n p_{k,j} x_{t,k} \quad (3.11)$$

donde $p_{k,j}$ es la probabilidad cambiar de estado, o permanecer en el mismo, además

$$\sum_{j=1}^n p_{k,j} = 1 \quad (3.12)$$

para $j, k = 1, 2, \dots, n$.

Y con $p_{k,j} \geq 0$ para $j, k = 1, 2, \dots, n$.

Entonces, utilizando el criterio de máxima entropía desarrollado en el capítulo anterior tenemos que la estimación de la matriz de probabilidades de transición está dada por:

$$Max(ME) = - \sum_{k,j} p_{k,j} \ln(p_{k,j}) \quad (3.13)$$

sujeto a:

$$x_{t+1,j} = \sum_{k=1}^n p_{k,j} x_{t,k} \text{ y } \sum_{j=1}^n p_{k,j} = 1.$$

La solución para el problema ME es:

$$\hat{p}_{k,j} = \frac{\exp(-\sum_t x_{t,k} \hat{\lambda}_{t,j})}{\sum_j \exp(-\sum_t x_{t,j} \hat{\lambda}_{t,j})}$$

donde $\hat{\lambda}$ es el vector $n - dimensional$ de multiplicadores de *Lagrange* relacionado.

3.3.5. Matrices Potencia

En este método se supone que hay homogeneidad de tiempo.

Sea $P(t)$ la matriz de transición crediticia correspondiente al periodo t y

asumiendo que es una matriz anual. Entonces la matriz correspondiente a periodos posteriores está dada por:

$$P(t + 1) = P(t) * P(1) \quad (3.14)$$

Capítulo 4

Comparación de Matrices de Transición de Riesgo de Crédito

En este capítulo se presentan algunas aplicaciones del concepto de entropía. Además se realiza la estimación de matrices de transición crediticia. Se presentan los resultados de la comparación entre los diferentes métodos utilizados para la estimar dichas matrices mediante el cambio en entropía relativa para obtener el resultado buscado: *un criterio de selección de modelos de riesgo de crédito*.

4.1. Introducción

En el mundo de las calificaciones crediticias existen diferentes modelos de riesgo de crédito para el cálculo de las respectivas matrices de transición. Dado el impacto económico que representa una mala estimación de éstas, resulta de gran importancia tener un criterio de selección de modelos de riesgo de crédito para hacer la estimación de dichas matrices.

Con base en la comparación de las matrices de transición crediticias generadas usando algunos de los métodos expuestos en el capítulo anterior se propone un criterio de selección.

Utilizando el cambio en entropía relativa para llevar a cabo la comparación de las matrices, se tomará como criterio de selección a la minimización del cambio en entropía relativa. Es decir, se tomará como mejor método de es-

timación de matriz de transición crediticia aquel cuyo cambio en entropía relativa sea mínimo en un periodo de tiempo determinado.

4.2. Aplicaciones de la Entropía

En esta sección se presentan algunas aplicaciones de la entropía, principalmente en la teoría de la comunicación y en economía. También es útil en campos sociales y en las ciencias de la vida.

4.2.1. Entropía en la Física

En esta área el concepto de entropía es central en las descripciones de la Termodinámica e iguala el universo como una totalidad.

La entropía es la magnitud termodinámica que expresa el grado de desorden molecular de los cuerpos aislados y que caracteriza termodinámicamente el estado de un sistema.

La entropía mide el grado de azar o aleatoriedad de la situación y la tendencia de los sistemas físicos a hacerse cada vez menos organizados y cada más aleatorios.

4.2.2. La Entropía en la Teoría de la Comunicación

El matemático Claude Elwood Shannon, aplicó (1948) una teoría matemática de la comunicación que posteriormente fue llamada Teoría de la Información. Definió la información como el grado de libertad de la fuente de información para elegir entre elementos de un idioma a fin de componer un mensaje determinado.

Este trabajo, de importancia fundamental en problemas de Comunicación, abrió nuevas trayectorias de investigación en las matemáticas. La comunicación se manifiesta en una variedad de comportamientos, procesos, y tecnologías cuyo significado se transmiten o derivan de la información.

En teoría de la comunicación la entropía es importante. El hecho de que

la información se mida por la entropía es natural, si se piensa que, la información se asocia al grado de libertad de elección que se tiene al construir los mensajes. Dada una fuente de información, se puede decir que ésta situación está muy organizada y no se caracteriza por un elevado grado de azar o de elección, si la entropía es baja.

Teoría de la Información

El término información no se refiere tanto a lo que se dice, sino a lo que se podría decir. Es decir, la información es la medida de la libre elección del mensaje.

Cuando un receptor de mensajes se enfrenta a la situación de tener que elegir entre dos mensajes alternativos, se dice que la información asociada a esta situación es la unidad.

Es un error decir, que uno u otro de los mensajes independientes representa una unidad de información. El concepto de información se refiere, no a los mensajes individuales sino, a la totalidad.

Si tenemos una fuente de información que produce mensajes por selección sucesiva de símbolos discretos entonces tiene una probabilidad de elegir diversos símbolos en una etapa del proceso que depende de las elecciones anteriores.

Entonces, cuantitativamente la información tiene su equivalente en el concepto físico de entropía. Ésta se expresa en términos de las probabilidades implicadas. La relación que la determina, se expresa en términos del logaritmo de las probabilidades, de modo que se tiene una generalización natural de la medida logarítmica para los casos simples.

La entropía (información, o libertad de elección) de cierta fuente de información se puede comparar con el valor máximo que esta entropía podría alcanzar si estuviese sujeta sólo, a la condición de que la fuente continúe empleando los mismos símbolos. La relación existente entre la entropía real y la máxima, se llama entropía relativa de una fuente.

El complemento de la entropía relativa, es decir uno menos la entropía relati-

va, se llama redundancia y es la fracción del mensaje que no está determinada por la libre elección del emisor, sino por las reglas estadísticas aceptadas que gobiernan el uso de los símbolos en cuestión. Se llama redundancia porque esta fracción del mensaje coincide con lo que en sentido ordinario se entiende por redundancia o dicho de otra manera, esta fracción del mensaje es innecesaria (y por tanto repetitiva o redundante) en el sentido de que si faltara, en el mensaje, éste seguiría completo, o podría completarse.

Ahora supongase que tenemos un conjunto de n símbolos independientes, o un conjunto de n mensajes completos independientes, cuyas probabilidades de elección son $p_1, p_2, p_3, \dots, p_n$, entonces la expresión real de la información será:

$$H = \sum_{i=1}^n p_i \log(p_i) \quad (4.1)$$

Veamos con unos ejemplos los valores que toma ésta expresión.

Supongamos que sólo estamos eligiendo entre dos mensajes posibles, cuyas probabilidades son p_1 para el primero y $p_2 = 1 - p_1$ para el segundo. Entonces la entropía, alcanza el valor mayor, cuando los dos mensajes son igualmente probables, es decir, $p_1 = p_2 = \frac{1}{2}$.

En este caso se dispone de libertad total para elegir entre los dos mensajes. En cambio, cuando un mensaje es más probable que otro (por ejemplo $p_1 > p_2$) el valor de H disminuye. Y cuando un mensaje es muy probable ($p_1 \approx 1$ y $p_2 \approx 0$), el valor de H es muy cercano a cero.

En el caso de que una de las probabilidades sea la unidad (certeza) y la otra cero (imposibilidad), H alcanza el valor cero (no hay libertad de elección, no hay información).

Entonces podemos concluir que H alcanza su mayor valor cuando las dos probabilidades son iguales, es decir, cuando es completamente libre e imparcial en la elección; y se reduce a cero cuando la libertad de elección desaparece.

La situación descrita anteriormente corresponde al caso singular en que los mensajes sean dos. Si los mensajes posibles son n , en vez de dos, entonces la entropía es máxima cuando las probabilidades de las diversas elecciones son aproximadamente iguales y la entropía tiene el valor mayor cuando la

libertad posible es mayor que en las elecciones sin que se ejerza presión o influencia hacia algunas de ellas.

Supongamos que una elección tiene una probabilidad próxima a uno y que las otras tienen probabilidades próximas a cero. En este caso se está influenciando hacia una elección en particular, y por tanto se tiene poca entropía.

4.2.3. La Entropía en Economía

Las medidas de concentración ¹ sirven para poner de relieve el mayor o menor grado de igualdad en el reparto del total de los valores de una variable económica. Por esta razón, son indicadores del grado de equidistribución de la variable. Desde el punto de vista de la equidad, otras medidas de posición o dispersión, no facilitan información relevante respecto al orden o igualdad de la variable que se estudia.

Las medidas de concentración se aplican especialmente a variables económicas por ejemplo los salarios de una empresa, las rentas que perciben las unidades económicas, el financiamiento que reciben los distintos municipios o estados de un país, la evolución de los ingresos realmente recibidos a lo largo del tiempo, la concentración industrial, etc..

Definiendo la concentración como la mayor o menor equidad en el reparto de la suma total de la variable considerada.

Entonces, dada una distribución de frecuencias de la variable x_i para $i = \{1, 2, \dots, n\}$, lo importante es saber hasta que punto :

$$\sum_{i=1}^n x_i = X \quad (4.2)$$

Está equitativamente distribuida.

Además de los estudios tradicionales por medio del índice de concentración de *Gini*, *Enri Theil* basándose en el concepto de entropía observó que el grado de desigualdad podría ser estudiado por medio de dicha función ya que la cuantía de cada variable podía expresarse como un cociente respecto del

¹Martín-Guzmán, M.P. y Martín Pliego, F.J.: "Curso básico de estadística económica". 30 edición. Editorial AC, pag.: 91 y sig.

total.

Entonces, tenemos:

$x_i \geq 0$ para todo $i = 1, 2, \dots, n$.

y

$$\sum_{i=1}^n \frac{x_i}{X} = 1.$$

Podiendo de esta manera razonar en términos de probabilidad.

Existen otras medidas que habitualmente utilizan los economistas, y también los sociólogos, en el estudio de la desigualdad en el reparto de una variable tales como el coeficiente de variación de *Pearson (CV)*, la *varianza del logaritmo natural (VLN)* y el índice de *Atkinson*.

La ventaja de utilizar una u otras medidas se ve justificada por las características de la distribución de frecuencias que aconsejan una u otra medida².

H. Theil afirma que el concepto de entropía puede ser aplicado a distintos problemas entre los que se encuentran los siguientes:

1. La medida de la desigualdad de los ingresos.
2. La concentración industrial.
3. La concentración del comercio internacional.
4. La evaluación de las encuestas.
5. Para ajustar relaciones de comportamiento.
6. Para el análisis de agregación en los modelos *input-output*.
7. Para los problemas de números índices de precios y cantidades.

Sea X una variable aleatoria discreta que toma los valores x_i , para $i = 1, 2, \dots, n$. Entonces la probabilidad de que ocurra cada uno de los sucesos vienen dadas por $P(x_i)$ y donde:

²Figart, D.M. (Richard Stockton College) y Lapidus, J. (Roosevelt University): "The impact of Comparable Worth on Earnings Inequality". *Work and Occupations*, vol.: 23, número 3, August 1996, pags.: 297-318.

$0 \leq P(x_i) \leq 1$ para toda $i = 1, 2, \dots, n$, y $\sum_{i=1}^n P(x_i) = 1$

Definiendo ahora entropía de la variable aleatoria X :

$$H(X) = \sum_{i=1}^n P(x_i) \log_2(P(x_i)) \quad (4.3)$$

Al ser la función de entropía una función logarítmica, está de acuerdo con el hecho de que, cuando un evento es muy poco probable y ocurre, su valor es mayor, que cuando ocurre un evento, que se consideraba que iba a suceder.

Como logaritmo de cero es infinito, por lo que cuanto más se aproxima, la probabilidad de ocurrencia de un evento a cero mayor será su logaritmo. Además la suma de los n valores están ponderadas por el peso que supone la probabilidad de ocurrencia de cada suceso.

En la definición de entropía intervienen los n productos de la probabilidad de que ocurra el suceso; $P(x_i)$ y el $\log_2 P(x_i)$. A este último término se le denomina autoinformación. Y es tanto mayor cuanto menor es la probabilidad de que el evento x_i ocurra.

El producto de n términos formados por el producto de $P(x_i)$ y el $\log_2 P(x_i)$ representan el valor medio de las informaciones que pueden proporcionar los resultados posibles de las variables aleatorias. La entropía, de la información, representa una medida de la incertidumbre.

Es pues la entropía una medida apropiada de medida de la desigualdad en el reparto de las variables económicas.

4.3. Generación de Matrices de Transición Crediticias

4.3.1. La Motivación

Supongamos que estamos en un escenario hipotético, en el cual, el espacio de estados es $E = \{1, 2, 3, 4\}$ y sabemos que una variable aleatoria X se distribuye uniformemente en E . Entonces, sabemos que la matriz de transición

92 Comparación de Matrices de Transición de Riesgo de Crédito

correspondiente debería ser:

$$U = \begin{pmatrix} 0,25 & 0,25 & 0,25 & 0,25 \\ 0,25 & 0,25 & 0,25 & 0,25 \\ 0,25 & 0,25 & 0,25 & 0,25 \\ 0,25 & 0,25 & 0,25 & 0,25 \end{pmatrix}.$$

Ahora, supongamos que para obtener dicha matriz de transición se generan números aleatorios para una variable aleatoria X , tal que $X \sim U(0, 1)$.

Por medio de una simulación en *Matlab* (ver apéndice A) se obtienen matrices de transición correspondientes a la cantidad de números utilizados en la simulación.

Usando la simulación para generar una matriz de transición con 100 números aleatorios, se obtiene la matriz:

$$U_{100} = \begin{pmatrix} 0,3 & 0,15 & 0,35 & 0,2003 \\ 0,1481 & 0,2222 & 0,2222 & 0,4077 \\ 0,1538 & 0,5 & 0,1538 & 0,1926 \\ 0,2308 & 0,1923 & 0,3462 & 0,2311 \end{pmatrix}.$$

Haciendo lo mismo pero con una simulación de 1000 números obtenemos la matriz siguiente:

$$U_{1000} = \begin{pmatrix} 0,1928 & 0,2530 & 0,2691 & 0,2851 \\ 0,2705 & 0,2500 & 0,2336 & 0,2459 \\ 0,2672 & 0,2026 & 0,2069 & 0,3233 \\ 0,2628 & 0,2664 & 0,2190 & 0,2518 \end{pmatrix}.$$

Y generando 10000 números aleatorios obtenemos la siguiente matriz:

$$U_{10000} = \begin{pmatrix} 0,2465 & 0,2441 & 0,2530 & 0,2563 \\ 0,2523 & 0,2370 & 0,2646 & 0,2461 \\ 0,2445 & 0,2374 & 0,2581 & 0,2600 \\ 0,2419 & 0,2514 & 0,2526 & 0,2541 \end{pmatrix}.$$

Claramente si aumenta la cantidad de números aleatorios generados para crear la correspondiente matriz de transición, ésta debe tender hacia la matriz U .

Calculando el cambio en entropía relativa entre cada una de las matrices generadas y la matriz original U tenemos que:

1. Entropía Relativa entre U y $U_{100} = \sum_{m=1}^{100} x_m \ln\left(\frac{x_m}{u_m}\right) = 0.494492903$
2. Entropía Relativa entre U y $U_{1000} = \sum_{m=1}^{1000} x_m \ln\left(\frac{x_m}{u_m}\right) = 0.033621832$
3. Entropía Relativa entre U y $U_{10000} = \sum_{m=1}^{10000} x_m \ln\left(\frac{x_m}{u_m}\right) = 0.001789474$

Entonces conforme la matriz de transición correspondiente se acerca a la real, la medida de entropía relativa va decreciendo.

Ahora tomamos una variable aleatoria Y que se distribuya Normal con parámetros $(0, 1)$, es decir $Y \sim N(0, 1)$ en el mismo escenario que X .

Una vez más, haciendo una simulación en *Matlab* y con un código (ver apéndice B) análogo (haciendo los cambios adecuados) al de la distribución uniforme obtenemos las siguientes matrices de transición.

Primero, generando 100 números aleatorios obtenemos la siguiente matriz:

$$N_{100} = \begin{pmatrix} 0,0714 & 0,3571 & 0,4286 & 0,1429 \\ 0,1667 & 0,2333 & 0,5000 & 0,1000 \\ 0,1190 & 0,2857 & 0,4286 & 0,1667 \\ 0,2308 & 0,4615 & 0,2308 & 0,0769 \end{pmatrix}.$$

Ahora usando una simulación de 1000 números aleatorios la correspondiente matriz de transición es:

$$N_{1000} = \begin{pmatrix} 0,1404 & 0,3567 & 0,3626 & 0,1404 \\ 0,1948 & 0,3343 & 0,3605 & 0,1105 \\ 0,1705 & 0,3526 & 0,3266 & 0,1503 \\ 0,1449 & 0,3333 & 0,3478 & 0,1739 \end{pmatrix}.$$

Finalmente con una simulación de 10000 números aleatorios obtenemos la matriz:

$$N_{10000} = \begin{pmatrix} 0,1767 & 0,3129 & 0,3420 & 0,1685 \\ 0,1615 & 0,3395 & 0,3433 & 0,1556 \\ 0,1534 & 0,3515 & 0,3323 & 0,1628 \\ 0,1412 & 0,3427 & 0,3660 & 0,1500 \end{pmatrix}.$$

Nuevamente, calculando la medida de entropía entre estas matrices de transición y la matriz de transición real (U), tenemos que:

1. Entropía Relativa entre U y $N_{100} = \sum_{i=1}^{100} x_m \ln(\frac{x_m}{u_m}) = 0.701422644$
2. Entropía Relativa entre U y $N_{1000} = \sum_{i=1}^{1000} x_m \ln(\frac{x_m}{u_m}) = 0.325739313$
3. Entropía Relativa entre U y $N_{10000} = \sum_{i=1}^{10000} x_m \ln(\frac{x_m}{u_m}) = 0.27856005$

Entonces la medida de entropía de las matrices de transición generadas con la cantidad de números aleatorios correspondiente y con la distribución Uniforme es menor que la de las matrices generadas por la distribución Normal.

Además se tiene un cambio en entropía relativa promedio para cada simulación de:

	Cambio en Entropía Relativa
Uniforme	0.212152609
Normal	0.257070274

Podemos concluir entonces que si se utiliza un modelo adecuado para estimar matrices de transición la medida de entropía se minimiza.

Motivados en estos resultados se utilizará el cambio en entropía relativa como discriminante para las matrices de transición de los modelos de riesgo de crédito.

4.4. Comparación de Matrices de Transición de Riesgo Crediticio con Entropía Relativa

4.4.1. Minimizando la Entropía Relativa

En el caso donde tenemos una matriz de referencia X^{t-1} el problema de escoger la mejor estimación se resuelve minimizando la distancia $D(X^{t-1}, Y^t)$.

Entonces tomando como distancia entre matrices de transición a la entropía relativa

$$E_r(Y) = \sum_{i,j} y_{i,j} \ln\left(\frac{y_{i,j}}{x_{i,j}}\right) \quad (4.4)$$

De esta forma, midiendo el grado de similitud entre dos matrices con la entropía relativa, la mejor estimación de matriz de transición, es la que minimice la entropía relativa.

$$\min E_r(Y)$$

sujeito a:

$$\sum_j^n y_{i,j} = 1$$

$$y_{i,j} \geq 0$$

4.4.2. Aplicación

En esta sección se generaran matrices de transición crediticias para tres periodos posteriores al de la matriz original.

Para generar dichas matrices serán utilizados los métodos *OM*, *OMG* y suponiendo homogeneidad el de *Matrices Potencia*.

Una vez obtenidas las respectivas matrices se realizará la comparación con el cambio en entropía relativa para tener conocimiento de la mejor estimación.

Dada la siguiente matriz de transición de calificaciones de riesgo de crédito anual correspondiente al periodo t :

Matriz de Transición de Calificaciones Crediticias Original para el periodo t

	AAA	AA	A	BBB	BB	B	C	D
AAA	0.88658	0.10294	0.01017	0	0.00031	0	0	0
AA	0.01079	0.88705	0.09553	0.00342	0.00145	0.00145	0	0.00031
A	0.00063	0.02876	0.90205	0.05919	0.0074	0.00177	0.0001	0.0001
BBB	0.00053	0.00339	0.07069	0.85237	0.06053	0.01005	0.00085	0.00159
BB	0.00033	0.00077	0.00557	0.0568	0.83571	0.08083	0.00535	0.01464
B	0.00011	0.00044	0.00174	0.00652	0.06595	0.82702	0.0276	0.07062
C	0	0	0.0066	0.0105	0.0305	0.0611	0.6297	0.2616
D	0	0	0	0	0	0	0	1

Utilizando el método *OM* (desarrollado en el capítulo anterior) estimamos la matriz de transición para el mismo periodo y de igual manera estimando para periodos posteriores (Ver apéndice C).

Ahora, utilizando el método *OMG* (desarrollado en el capítulo anterior) estimamos matrices de transición de riesgo de crédito para los mismos periodos

que con el método *OM* (Ver apéndice D).

Por último, con base a el método de *Matrices Potencia* calculamos las matrices de transición correspondiente a los mismos periodos de tiempo que los métodos anteriores (ver apéndice E).

Ahora comparando las matrices de transición obtenidas con los métodos *OM* y *OMG*, basados en el cambio de entropías relativa tenemos:

Entropías Relativas

	<i>OM</i>	<i>OMG</i>	<i>Matr. Pot.</i>
<i>t</i>	1.71793E-06	0.00041766	
<i>t + 1</i>	0.380922652	0.382749995	0.380499571
<i>t + 2</i>	0.182919547	0.183243248	0.182824736
<i>t + 3</i>	0.108978355	0.109927848	0.109404409

Como se puede observar en la tabla anterior el método *OM* es el de menor cambio en entropía relativa. Entonces con base al criterio de selección expuesto el método *OM* da una mejor estimación de matrices de transición que los otros dos.

De esta forma se pudo llegar a una selección de modelos de riesgo de crédito basándose en el cambio en la entropía relativa y con ello lograr minimizar el riesgo de crédito implícito en una mala estimación de las correspondientes matrices de transición crediticia.

El criterio de escoger el modelo que minimiza la entropía relativa entre dos periodos es consistente con el así llamado, principio de mínima producción de entropía para sistemas termodinámicos fuera del equilibrio.[10]

Capítulo 5

Conclusiones

Después del trabajo realizado se puede concluir diciendo que aplicando los conceptos de entropía para comparar diferentes modelos de estimación del riesgo de crédito es posible obtener un criterio de discriminación de modelos de riesgo de crédito.

Es importante mencionar que con un criterio así se puede llegar a acotar las pérdidas económicas originadas por el riesgo de crédito. Además de poder tomar decisiones en cuanto a adoptar un modelo para estimar la probabilidad de incumplimiento.

Una observación importante es el hecho de que basados en el principio de máxima entropía se puede llegar a generar la correspondiente matriz de transición crediticia mediante métodos numéricos. Entonces con los resultados obtenidos y utilizando el criterio de mínima entropía relativa resulta de gran interés comparar este modelo con los otros.

Por otra parte, las analogías con sistemas físicos termodinámicos fuera del equilibrio, ejemplificada por el principio de mínima producción de entropía pueden ser muy útiles en el estudio de problemas financieros.

Como se presenta en el último capítulo el concepto de entropía tiene diversas aplicaciones no sólo en el mundo financiero sino en diversas ramas. Tomando como punto de partida el trabajo aquí realizado se puede tomar el concepto de entropía como una muy buena alternativa para trabajar dife-

rentes aplicaciones financieras.

Apéndice A

Simulación Uniforme

El código en *Matlab* para generar las matrices con **distribución uniforme**.

```
function U=uniforme(n)
    X=unifrnd(0,1,n,1);
    A=zeros(4,1);
    for i=1:length(X)
        if X(i) <= .25
            A(i)=1;
        elseif X(i) <= .5
            A(i)=2;
        elseif X(i) <= .75
            A(i)=3;
        else
            A(i)=4;
        end
    end
    P=zeros(4,4);
    for i=1:length(A)-1;
        if A(i)==1
            if A(i+1)==1
                P(A(i),1)=P(A(i),1)+1;
            elseif A(i+1)==2
                P(A(i),2)=P(A(i),2)+1;
            elseif A(i+1)==3
                P(A(i),3)=P(A(i),3)+1;
            end
        end
    end
end
```

```
elseif A(i+1)==4
P(A(i),4)=P(A(i),4)+1;
end
end
if A(i)==2
if A(i+1)==1
P(A(i),1)=P(A(i),1)+1;
elseif A(i+1)==2
P(A(i),2)=P(A(i),2)+1;
elseif A(i+1)==3
P(A(i),3)=P(A(i),3)+1;
elseif A(i+1)==4
P(A(i),4)=P(A(i),4)+1;
end
end
if A(i)==3
if A(i+1)==1
P(A(i),1)=P(A(i),1)+1;
elseif A(i+1)==2
P(A(i),2)=P(A(i),2)+1;
elseif A(i+1)==3
P(A(i),3)=P(A(i),3)+1;
elseif A(i+1)==4
P(A(i),4)=P(A(i),4)+1;
end
end
if A(i)==4
if A(i+1)==1
P(A(i),1)=P(A(i),1)+1;
elseif A(i+1)==2
P(A(i),2)=P(A(i),2)+1;
elseif A(i+1)==3
P(A(i),3)=P(A(i),3)+1;
elseif A(i+1)==4
P(A(i),4)=P(A(i),4)+1;
end
end
end
```

```
for i=1:4
P(i,:)=P(i,+)/sum(P(i,:))
end
U=P;
```

Apéndice B

Simulación Normal

El código en *Matlab* para generar las matrices con distribución normal

```
function M=normal(n)
X=RANDOM('norm',0,1,n,1);
A=zeros(4,1);
for i=1:length(X) if X(i) <= -1
A(i)=1;
elseif X(i) <= 0
A(i)=2;
elseif X(i) <= 1
A(i)=3;
else
A(i)=4;
end
end
P=zeros(4,4);
for i=1:length(A)-1;
if A(i)==1
if A(i+1)==1
P(A(i),1)=P(A(i),1)+1;
elseif A(i+1)==2
P(A(i),2)=P(A(i),2)+1;
elseif A(i+1)==3
P(A(i),3)=P(A(i),3)+1;
```

```
elseif A(i+1)==4
P(A(i),4)=P(A(i),4)+1;
end
end
if A(i)==2
if A(i+1)==1
P(A(i),1)=P(A(i),1)+1;
elseif A(i+1)==2
P(A(i),2)=P(A(i),2)+1;
elseif A(i+1)==3
P(A(i),3)=P(A(i),3)+1;
elseif A(i+1)==4
P(A(i),4)=P(A(i),4)+1;
end
end
if A(i)==3
if A(i+1)==1
P(A(i),1)=P(A(i),1)+1;
elseif A(i+1)==2
P(A(i),2)=P(A(i),2)+1;
elseif A(i+1)==3
P(A(i),3)=P(A(i),3)+1;
elseif A(i+1)==4
P(A(i),4)=P(A(i),4)+1;
end
end
if A(i)==4
if A(i+1)==1
P(A(i),1)=P(A(i),1)+1;
elseif A(i+1)==2
P(A(i),2)=P(A(i),2)+1;
elseif A(i+1)==3
P(A(i),3)=P(A(i),3)+1;
elseif A(i+1)==4
P(A(i),4)=P(A(i),4)+1;
end
end
end
```

```
for i=1:4  
P(i,:)=P(i,+)/sum(P(i,:))  
end  
M=P;
```


Bibliografía

- [1] Alexander Kreinin y Marina Sidelnikova, *Regularization Algorithms for Transition Matrices*, Algo Research Quartely, vol. 4, Marzo 2001.
- [2] Alfred Hamerle y daniel Rosch, *Parametrizing Credit Risk Models*, artículo preliminar.
- [3] Amos Golan, George Judge, Douglas Miller, *Maximun Entropy econometrics: Robust Estimation with Limited Data*, edit. John Wiley and Sons, 1996.
- [4] Ana Bertha Palacios Paz, *Modelos de Riesgo de Crédito*, Tesis para Licenciatura de Actuaría, Facultad de Ciencias, UNAM, 2000.
- [5] Anatoliy Antonov y Yanka Yanakieva, *Transition Matrix Generation*, International Conference on Computer Systems and Technologies, 2004.
- [6] Boletines de Comite de Basilea II (Riesgo de crédito) (1995).
- [7] Dennis Glennon y Amos Golan, *A Markov Model of Bank Failure Estimated Using an Information-Theoretic Approach*, artículo preliminar, Marzo 2003.
- [8] Gabriela Conde, Fabio Malacrida y Ricardo Selves, *Valuación de Instrumentos sujetos a Riesgo de Crédito*, artículo preliminar, Julio 2003.
- [9] Guillermo López Dumrauf, *Rentabilidad y Riesgos de Activos financieros*, artículo preliminar, Abril 2004.
- [10] Ilya Prigogine, *Introduction a la thermodynamique des processus irréversibles*, edit Gallimard. 1996.

- [11] J.N. Kapur., *Maximun Entropy Models in Science and Engineering*, edit John Wiley and Sons, 1989.
- [12] Javier Márquez Diez-Canedo, *Suficiencia de Capital y Riesgo de Crédito en Carteras de Préstamos Bancarios*, Dirección General de Análisis de Sistema Financiero, Banco de México, abril 2002.
- [13] José Luis Cuenca Tadeo *Tres Lugares Comunes para la Entropía*, artículo preliminar.
- [14] Joseph Stampfli y Victor Goodman *The Mathematics of Finance, Modeling and Hedging*, edit Thomson, 2002.
- [15] Mario Zambrano Berendsohn *Un Modelo Crediticio Básico, Regulación Prudencial, Volatilidad Cambiaria y Medición de Riesgos*, artículo preliminar, Agosto 2004.
- [16] Martha Galicia Romero *Nuevos Enfoques de Riesgo de Crédito*, artículo preliminar, 2003.
- [17] Massaaki Kijima, *Stochastic Processes with Applications to Finance*, edit. Chapman and Hall, 2003.
- [18] U. Narayan Bhat y Gregory K. Miller, *Elements of Applied Stochastic Processes*, tercera edición. edit. Wiley, 2002.
- [19] Uwe Blien y Friedrich Graef, *Entropy Optimizaing Methods for the Estimation of Tables*, 21st Annual Conference of the Gesellschaft für Klassifikation, University of Potsdam, 1997.
- [20] www.Gacetafinanciera.com.
- [21] www.moody.com.
- [22] www.standardandpoors.com.
- [23] www.Sudaval.com.
- [24] Yusult Jafry y Til Schuermann, *Measurement, Estimation and Comparison of Credit Migration Matrices*, artículo preliminar, Marzo 2004.