

Universidad EAFIT

Práctica Investigativa 3: construcción de un indicador de estado financiero

Grupo de Investigación de Simulación y Modelación
Matemática Tutora: Myladis Cogollo

Laura Carolina Cardona Mesa
Estudiante Ingeniería Matemática
Noviembre de 2009

Introducción

En una entidad financiera se tiene interés por conocer el estado financiero de sus clientes. En especial, cuando los clientes son empresas se hace necesario monitorear los indicadores financieros de las mismas para establecer cuándo éstos se empiezan a deteriorar y generar alertas sobre dicho cliente pues puede estar en dificultades para, por ejemplo, cumplir sus obligaciones con la entidad. En este caso es relevante entonces realizar un análisis del estado financiero de las compañías para determinar su estado respecto a las compañías del sector y posteriormente utilizar esta información para alimentar un modelo de riesgo de crédito en el área de auditoría de la entidad financiera.

En este trabajo se tiene como objetivo la construcción de un indicador que, partiendo de la información de múltiples variables de estado financiero de una compañía, indique el estado general de la misma con respecto a su sector. Con este objetivo además se busca hacer uso de herramientas estadísticas para tecnificar un proceso que hasta ahora se hace manualmente y a juicio del experto.

Hasta el momento este proceso se realiza tomando la información de indicadores financieros de empresas y evaluando el comportamiento del indicador de la empresa con respecto al promedio del sector, además se evalúa la probabilidad de que el indicador se incremente y se decide con base en lo anterior si el indicador tiene un comportamiento alarmante (deterioro). Este proceso se repite para varios indicadores y se da un concepto de alarma final, ponderando las alarmas individuales de cada indicador según el juicio del experto. Lo que se busca con este trabajo es establecer la ponderación de las variables de manera más técnica y elegir las variables más relevantes para cada sector, teniendo en cuenta que sólo se tiene información sobre indicadores financieros más no existe ninguna variable de respuesta con la que se puedan comparar estas variables de entrada.

Para esto se utilizó la técnica estadística de componentes principales, logrando obtener un indicador que basado en alertas individuales genera un alerta global, utilizando ponderaciones dadas por la técnica elegida. De esta manera se obtiene un indicador que puede ser replicado y usado para todas las empresas del sector elegido y que genera un estándar de calificación para el sector. Además, la metodología descrita permite replicar el mismo proceso para empresas de diferentes sectores.

Marco teórico

Según las características del problema mencionadas anteriormente se buscaron herramientas estadísticas que permitieran abordar el problema, teniendo en cuenta que en este caso, la principal característica es que se deben buscar metodologías que no requieran variables de salida u objetivo debido a que se dispone de información de muchas variables cuantitativas (los indicadores financieros de cada empresa) pero no se cuenta con una o algunas variables de respuesta, es decir, variables que sean el resultado directo o indirecto de las variables de entrada. Esta característica del problema limita el número de técnicas estadísticas que se pueden usar para alcanzar el objetivo propuesto. Entre ellas se tienen técnicas de estadística multivariada como análisis de componentes principales, análisis de correspondencias, análisis cluster, entre otras. Según las características del problema se opta por estudiar y usar el análisis de componentes principales.

Análisis de componentes principales

En ocasiones si se toman muchas variables sobre un fenómeno lo normal es que estén relacionadas o que midan lo mismo bajo distintos puntos de vista es por esto que se hace necesario reducir el número de variables pero sin perder información. Esta metodología fue desarrollada por Pearson a finales de 1800 y estudiada por Hotelling en los años 30 (Shlens, 2005).

El Análisis de componentes principales busca reducir un conjunto de n variables correlacionadas en $k < n$ variables no correlacionadas, llamadas componentes principales. Las nuevas variables son combinaciones lineales de las originales y son ortogonales entre sí. Éstas buscan capturar la mayor cantidad de varianza disponible en la totalidad de las variables. Con esto se busca que el número de componentes resultantes sea pequeño, que la solución obtenida sea interpretable y que se pierda la menor cantidad de información (Marín, 2006).

Uno de los métodos para el cálculo de componentes principales es mediante la matriz de correlación o covarianza de las variables originales. De esta manera, los componentes son los vectores propios de la matriz y se eligen los componentes principales con valor propio mayor a 1 como los componentes que será relevantes (Smith, 2002). El determinante de la matriz debe ser cercano a cero lo que indica que de hecho existe correlación entre las variables, si este hecho no se da, el uso de componentes principales sería inadecuado. Para verificar este hecho se usa la prueba de esfericidad de Barlett (SPSS Inc., 2007).

Adicional a esto, se debe verificar la medida de adecuación muestral, que indica la magnitud de la relación lineal entre las variables. Para esto se usa la prueba de adecuación muestral de Kaiser-Meyer-Olkin (KMO) (Smith, 2002).

En el análisis de componentes principales es usual utilizar la herramienta gráfica *biplot*. Un *biplot* es un gráfico de dispersión de los componentes principales y líneas por cada una de las variables presentes en el componente con el fin de mostrar cómo cada una de las variables contribuye a éste. Las líneas intersecan en el origen y representan las variables originales (Statgraphics, 1999).

Metodología

Se tiene información de Información de 28 indicadores financieros de 620 empresas (aproximadamente 2000 registros) para un sector económico piloto. Estas variables se normalizaron pues existían unidades de medida diferentes y en diferentes escalas, lo que podría alterar los resultados de los componentes principales. Con esta información preparada, se utilizaron los paquetes estadísticos *SPSS* (SPSS Inc., 2008) y *Statgraphics* (Statgraphics, 1999) para aplicar el análisis de componentes principales utilizando la matriz de correlación y su determinante, Prueba de esfericidad de Barlett y la prueba de adecuación muestral KMO como parámetros de verificación del procedimiento (Balbuena y Casas, 2003).

Inicialmente se aplicó el análisis de componentes principales a las 28 variables, buscan en los resultados del análisis identificar grupos de variables que formaran cada componente y verificando los parámetros descritos anteriormente. En esta primera etapa también se buscó identificar variables que fueran candidatas a eliminar. Posteriormente se realizó análisis de componentes principales por separado a cada uno de los grupos identificados en la primera etapa buscando verificar si conformaban un solo componente o si al ser evaluados entre menos variables a parecía nuevamente más componentes. Una vez se tuvieron definidos los grupos principales se buscó incluir variables que quedaron excluidas o en grupos más pequeño en los grupos más grandes, intentado obtener menor número de componentes. Durante el proceso, se verificó con el experto la posibilidad de eliminar algunas variables que tenía pobre desempeño dentro de los componentes llegando finalmente a 18 variables de las cuales 13 están agrupadas en 5 componentes y las restantes se incluyen individualmente utilizando criterios de alarma específicos.

Adicional a esto, se utilizó la herramienta *biplot* para evaluar gráficamente la disposición de cada variable dentro del componente. Como se puede ver en la figura 1, cada gráfico representa cada uno de los cinco componentes. Las líneas azules representan cada uno de los indicadores financieros presentes en el componente. La longitud de la línea indica qué tan importante es el indicador financiero dentro del componente. El eje x es el componente elegido y el eje y es otro componentes cualquiera. Con esto se ve que en todos los casos las variables agrupadas apuntan en una dirección similar y están más cercanas al componente 1 que al 2. Esto nos muestra que la agrupación elegida es correcta para cada caso y que las variables realmente se relacionan y quedan resumidas correctamente por el componente.

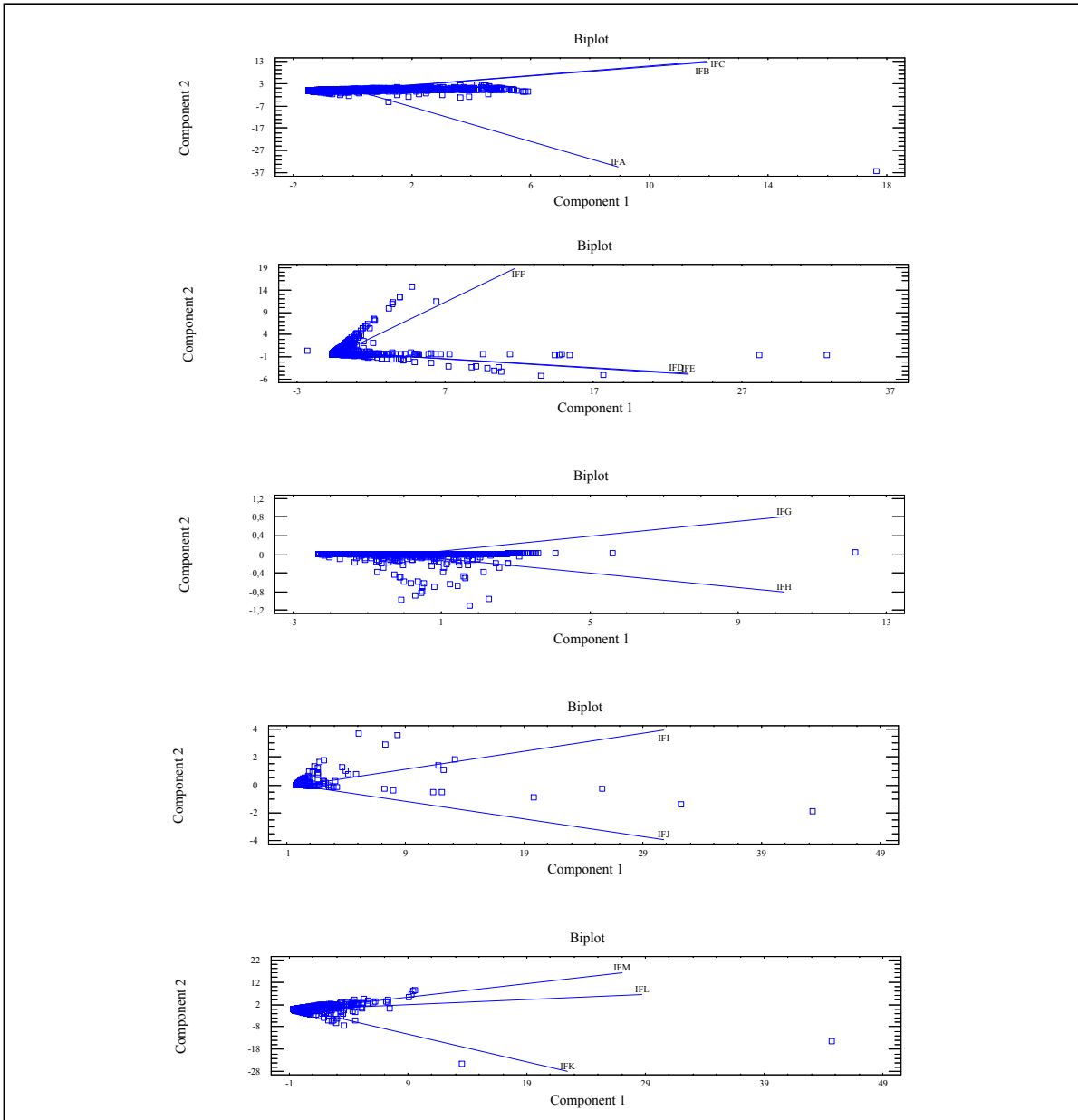


Figura 1: gráficos *bipot* para los 5 componentes creados

Resultados

Luego de realizar el proceso anteriormente descrito se obtuvieron cinco componentes. Cada componente está determinado por las siguientes expresiones, obtenidas con el paquete estadístico utilizado:

$$\text{componente1} = 0,303IF_A + 0,405IF_B + 0,405IF_C$$

$$\text{componente2} = 0,457IF_D + 0,457IF_E + 0,227IF_F$$

$$\text{componente3} = 0,501IF_G + 0,501IF_H$$

$$\text{componente4} = 0,503IF_I + 0,503IF_J$$

$$\text{componente5} = 0,440IF_K + 0,467IF_L + 0,365IF_M$$

En estas ecuaciones los valores de los indicadores financieros IF_x , $x = A, B, \dots, M$, son valores normalizados, es decir,

$$IF_x = \left(\frac{if_x - \overline{if_x}}{S_{if_x}} \right), \quad X = A, B, \dots, M,$$

donde if_x es el valor del indicador financiero, $\overline{if_x}$ es la media muestral del indicador y S_{if_x} es la varianza muestral de la variable.

De esta manera cada una de las 13 variables está presente en uno solo de los 5 componentes y recibe una ponderación dentro de ese componente según su importancia en dentro de éste. Con esto se logra obtener las ponderaciones esperadas y aunque no es una ponderación general, es una ponderación dentro de grupos de variables similares.

Cada uno de los registros que se tiene en los datos disponibles recibió una calificación de cada componente, según sus valores de cada indicador financiero. Esta calificación es proporcionada automáticamente por el paquete estadístico utilizado.

Para evaluar el comportamiento de cada individuo respecto al sector se deben establecer valores ideales de cada variable, para luego calificar a cada registro según este parámetro. Para esto se considera el valor ideal que tomaría el componente si se tuviera una empresa en condiciones perfectas y luego el valor de cada empresa se analiza con respecto a éste valor ideal. Por ejemplo, para el componente 1 se tiene:

$$\text{componente1} = 0,303IF_A + 0,405IF_B + 0,405IF_C$$

Pero como el valor del indicador financiero IF_x es un valor normalizado, entonces se tiene realmente:

$$componente1 = 0,303 \left(\frac{if_A - \overline{if_A}}{S_{if_A}} \right) + 0,405 \left(\frac{if_B - \overline{if_B}}{S_{if_B}} \right) + 0,405 \left(\frac{if_C - \overline{if_C}}{S_{if_C}} \right),$$

donde if_x es el valor del indicador financiero, $\overline{if_x}$ es la media muestral del indicador y S_{if_x} es la varianza muestral de la variable. Con esto tenemos que el valor máximo del componente se puede escribir como:

Valor máximo comp1

$$= 0,807 \left(\frac{VIif_A - \overline{if_A}}{S_{if_A}} \right) + 0,857 \left(\frac{VIif_B - \overline{if_B}}{S_{if_B}} \right) + 0,669 \left(\frac{VIif_C - \overline{if_C}}{S_{if_C}} \right)$$

Donde $VIif_x$ es el valor ideal del indicador financiero. En este caso, según la naturaleza de los indicadores financieros no se puede establecer un valor ideal pues estos indicadores no se encuentran definidos dentro de un rango. Por esta razón se estableció que el valor ideal de cada indicador financiero es:

$$VIif_x = \max(if_x)_s + 0.5S_{if_x}$$

para indicadores que son positivos, es decir aquéllos en los que el estado de la empresa es mejor entre mayor sea el indicador y,

$$VIif_x = \min(if_x)_s - 0.5S_{if_x}$$

para indicadores financieros que son negativos, es decir, aquéllos en los que el estado de la empresa es mejor entre menor sea el valor del indicador.

Con esto en cuenta, se crea un índice correspondiente para cada componente que busca crear valores entre cero y 100, donde 100 es el mejor comportamiento posible, de la siguiente manera:

$$\text{Índice } i = \frac{componentei}{\text{Valor máximo comp } i} * 100$$

Luego de tener este índice, se establecen los siguientes rangos de alerta:

Valor del índice	Alerta	
Entre 100 y 75	Sin alerta	0
Entre 75 y 50	Sin alerta	0
Entre 50 y 25	Alerta media	1
Entre 25 y 0	Alerta crítica	2

Con estos índices creados de la manera descrita se procede a evaluar cada individuo y a emitir una alerta por cada índice. Dado que existen 5 indicadores adicionales que no están en los componentes se crean alertas para cada uno de ellos. En este caso se crean tres alertas: dos alertas son resultado de la comparación entre dos indicadores y la tercera alerta es del comportamiento del indicador restante.

Posteriormente se crea una alerta total, comprendida por todas las alertas del registro, es decir, la sumatoria de los resultados de ocho alertas. Luego, dado que cada empresa puede tener varios registros o individuos, de diferentes años o meses, se crea una alerta por empresa definida como el promedio de las alertas totales de cada uno de los registros de la empresa. Esta alerta por empresa resulta entre los siguientes valores:

Valor Alerta Empresa	Alerta global
Entre 0 y 4	Sin alerta
Entre 4 y 8	Sin alerta
Entre 8 y 12	Alerta media
Entre 12 y 16	Alerta crítica

La definición de esta alerta global para cada empresa se tienen entonces lo que se busca: poder monitorear el estado financiero de una empresa según las demás empresas de su sector y emitir una alerta sobre ese estado financiero, buscando identificar el deterioro antes de que sea demasiado tarde.

Conclusiones

Se pudo estructurar una metodología para evaluar el comportamiento de una empresa respecto a las empresas del sector utilizando análisis de componentes principales y definiciones estadísticas sencillas. Esta metodología puede ser fácilmente replicada para evaluar nueva información de las empresas a medida que está disponible, nuevas empresas del sector y otros sectores, de una manera sencilla y fácil de entender.

La manera como está estructurada la metodología permite al experto que usa y evalúa el estado de las empresas detectar cuál es el componente o grupo de variables que se encuentra en estado más crítico e incluso se puede monitorear por componente, por registro o por empresa. Además se encuentra disponible en una hoja de cálculo lo que permite su fácil uso por parte de otras personas y modificación.

Finalmente, se pudo calcular ponderaciones de una manera más técnica, que era lo que se esperaba del trabajo.

Este tipo de trabajos tiene diversos aportes: dentro del grupo de investigaciones, permite apoyar el estudio de metodologías para construcción de indicadores; en Ingeniería Matemática, crea una relación entre la empresa y el pregrado (estudiantes y profesores); y en la empresa, propone soluciones que tecnifican desarrollos intuitivos y manuales.

Como trabajo futuro se plantea buscar una variable de respuesta que permita utilizar otras técnicas estadísticas para evaluar si con esto se obtienen mejores resultados. Incluso se podría utilizar la alerta global obtenida con este método como una nueva variable de respuesta u objetivo.

Bibliografía

Balbuena, C., J.R. Casas. Aplicación del análisis factorial a la valoración por parte de los estudiantes de las asignaturas de la etsIccp de Barcelona en sus distintas titulaciones. Actas del congreso internacional de Enseñanza en Ingeniería civil, Universidad Castilla La Mancha, Ciudad Real, 2003. <http://www-ma3.upc.edu/users/balbuena/Publications.html>

Marín, J.M. Notas de Clase. Análisis de Componentes Principales. Departamento de Estadística. Universidad Carlos III de Madrid. España. 2006. <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/AMult/tema3am.pdf>

Shlens, J. A Tutorial on Principal Component Analysis. Systems Neurobiology Laboratory. Salk Institute for Biological Studies. 2005. www.sn1.salk.edu/~shlens/pub/notes/pca.pdf

Smith, L.I. A tutorial on Principal Components Analysis. Cornell University. USA. 2002. <http://users.ecs.soton.ac.uk/hbr03r/pa037042.pdf>

SPSS Statistics 17.0, Versión 17.0.0, 2008. SPSS Inc.

SPSS Inc. Clementine® 12.0 Modeling Nodes. Copyright © 2007 by Integral Solutions Limited. ISBN-13: 978-1-56827-397-6. ISBN-10: 1-56827-397-5.

Statgraphics Plus for Windows 4.0. Copyright 1994-1999. Statistical Graphics Corp.