

1
00365



UNIVERSIDAD NACIONAL AUTONOMA ¹ DE MEXICO

POSGRADO EN CIENCIAS MATEMATICAS
FACULTAD DE CIENCIAS

UN MODELO MULTIVARIADO DE CLASIFICACION BORROSA PARA D3F

T E S I S
QUE PARA OBTENER EL GRADO ACADEMICO DE
MAESTRO EN CIENCIAS
(M A T E M A T I C A S)
P R E S E N T A :
FIS. ARTURO VAGUILAR VAZQUEZ

DIRECTOR DE TESIS:
DR. ONESIMO HERNANDEZ LERMA



FACULTAD DE CIENCIAS
UNAM

MEXICO, D. F.

MAYO 2003

TESIS CON
FALLA DE ORIGEN



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.

A LA FAMILIA AGUILAR PEÑA

***Porque, aunque no siempre se puede obtener todo lo que se desea,
cuando lo intentamos, podemos obtener lo que necesitamos.***

TESIS CON
FALLA DE ORIGEN

Agradezco al Dr. Onésimo Hernández Lerma todo el apoyo brindado, sin el cual no hubiera sido posible culminar apropiadamente este proyecto.

TESIS CON
FALLA DE ORIGEN

Prefacio

Vivimos en la era de la información. Cada día somos materialmente arrasados por la gran cantidad de información que nos llega desde las distintas fuentes de información disponibles: libros, revistas, periódicos, correos electrónicos, faxes, mensajes circulando en la red INTERNET, teléfonos celulares, etc. Se ha vuelto entonces una necesidad urgente poder disponer de herramientas que permitan navegar en este gran mar de información de manera que nos permita localizar o, al menos vislumbrar, rutas de acceso para desenmarañar la información disponible. También en este tenor, desde el punto de vista de la investigación científica, se requiere de herramientas que nos permitan la posibilidad de identificar, dado un conjunto de datos, las posibles estructuras que subyacen al interior del mismo. La búsqueda de estas estructuras subyacentes es, sin duda, uno de los objetivos fundamentales de la investigación científica.

El presente trabajo ofrece una alternativa en esta continua búsqueda de las estructuras ocultas dentro de grandes volúmenes de información a los que se pueden acceder hoy día. En particular se propone un modelo de clasificación borrosa adecuado para el estudio de los llamados datos de tres formas, o datos provenientes de estudios realizados a lo largo de cierto número de situaciones, que pueden ser temporales o no. El modelo propuesto requiere, para su aplicación, de un procedimiento algorítmico; el cual es desarrollado en detalle en un par de aplicaciones de índole económica.

El trabajo inicia en el capítulo 1 con una introducción a la teoría de los conjuntos borrosos, donde se establecen el concepto de conjunto borroso y las operaciones básicas asociadas a dichos conjuntos. Posteriormente, en el capítulo 2, se examinan los criterios de clasificación de datos, se definen los datos de 3 formas y los conceptos básicos del reconocimiento de patrones donde se inserta de manera natural el problema de la clasificación, proponiendo entonces el enfoque de la optimización de criterios múltiples para resolver el problema de clasificación de dicho tipo de datos. Finalmente, la estructura algorítmica del modelo propuesto es aplicada en el capítulo 3 para estudiar el patrón de evolución económica de México durante la segunda mitad del siglo XX.

Esperamos que la herramienta propuesta en este trabajo sea de utilidad en los problemas de clasificación similares a los aquí presentados.

TESIS CON
FALLA DE ORIGEN

Indice

- | | |
|--|-----------|
| 1. Introducción a la teoría de los conjuntos borrosos | 6 |
| 2. Un modelo multivariado de clasificación borrosa | 15 |
| 3. Aplicaciones del modelo de clasificación: El patrón de evolución económica de México en el contexto de la OCDE | 29 |

TESIS CON
FALLA DE ORIGEN

Capítulo 1. Introducción a la teoría de los conjuntos borrosos

TESIS CON
FALLA DE ORIGEN

1.1 La noción de borrosidad

Cuando requerimos de información y/o de conocimiento, el lenguaje común que empleamos generalmente está lleno de imprecisión y de vaguedad; es decir, de borrosidad. Por ejemplo, consideremos las siguientes afirmaciones:

- Un ingeniero industrial gana mucho dinero
- Las carreras sociales requieren de poco esfuerzo de los estudiantes
- La mayoría de la gente en México es pobre

En cada uno de los ejemplos anteriores es imposible definir con precisión nociones como "mucho dinero", "poco esfuerzo", o "pobre"; es decir, se trata de conceptos vagos o borrosos. Sin embargo, para nosotros estas afirmaciones no sólo tienen sentido, sino que, además, frecuentemente las empleamos para tomar decisiones al respecto.

Comparemos ahora esas afirmaciones con las siguientes: "ser divorciado", "tener más de 18 años de edad" o "pesar más de 78 kg", las cuales se definen con toda precisión y es fácil determinar si una persona u objeto satisface o no dichos enunciados. En contraste, cuando los enunciados son vagos o borrosos se vuelve un problema establecer con un sí o un no si las personas o los objetos satisfacen dichas afirmaciones, y más aún si se requiere determinar en qué magnitud las satisfacen.

Cuando nos piden respuestas categóricas "sí" o "no" es normal que pidamos información adicional sobre los conceptos involucrados; verbigracia: para afirmar que una nación N_1 es más pobre que otra N_2 primero debemos ser capaces de definir de una manera objetiva el conjunto de países pobres P , para luego clasificar a N_1 y N_2 respecto a P . En el caso de los enunciados borrosos empleados usualmente sucede que, no sólo es posible forzar una respuesta tipo "sí" o "no", sino que así se hace generalmente en la mayoría de los casos. Al ocurrir esto se pierde una cantidad de la información disponible al respecto, pues no se toma en cuenta el grado de pertenencia del concepto borroso propiamente dicho que poseen las personas o los objetos de estudio.

Es sabido también que en aquellos campos donde los enunciados empleados son precisos, ya sean ciertos o falsos (es decir, donde no se manejan enunciados imprecisos o borrosos), la metodología científica es empleada exitosamente. Sin embargo, en la actualidad el ideal de lograr producir máquinas que simulen el razonamiento humano, el cual se basa en el manejo de información imprecisa e incierta, se ha vuelto una piedra angular de la computación "suave", en contraposición con la computación "dura" convencional. En este esfuerzo se han vuelto esenciales la teoría de los conceptos borrosos y sus diversas aplicaciones, que hasta ahora se han desarrollado en los campos de la ingeniería y de las ciencias aplicadas. Los elementos matemáticos que forman la base de los conceptos borrosos existen desde mediados de los 60s, pero el surgimiento de aplicaciones exitosas como las técnicas de control borroso automático ha despertado el interés en el estudio de las bases matemáticas subyacentes a la teoría de los conceptos borrosos conocida como teoría de los conjuntos borrosos (TCB). En este contexto, el objetivo principal en general es utilizar la TCB para representar, manipular y obtener inferencias de este tipo de enunciados imprecisos o borrosos.

TESIS CON
FALLA DE ORIGEN

En el presente capítulo se discutirá la noción intrínseca de borrosidad en el lenguaje natural. Posteriormente se establecen las operaciones básicas para conjuntos borrosos y, finalmente, se comentan los diversos modelos matemáticos empleados para modelar la incertidumbre en los problemas del mundo real.

1.2 El concepto de conjunto borroso

El objeto matemático que emplearemos para representar conceptos borrosos es el concepto de conjunto borroso, el cual constituye el análogo del azar en la teoría de la probabilidad.

La modelación matemática de los conceptos borrosos fue introducida en 1965 por Lotfi Zadeh [Zadeh, 1965]. Hasta entonces se definía, usualmente, a un subconjunto A de un conjunto ordinario o conjunto rígido¹ dado U, por medio de la función indicadora o función característica, χ_A , la cual se define como:

$$\chi_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{cases} \quad (1.1)$$

Para cualquier $x \in U$, χ_A especifica si $x \in A$ o si $x \notin A$: de donde, en el enfoque ordinario o rígido, χ_A puede tomar sólo dos valores: 1 ó 0. Esta noción se generaliza en el enfoque borroso al permitir que las imágenes de los elementos pertenezcan al intervalo [0,1], en lugar de estar restringidos rigidamente a {0,1}, por medio de la siguiente definición.

Definición 1.1. Un subconjunto borroso del conjunto dado U es una función $A: U \rightarrow [0,1]$.

De acuerdo con la definición 1, aquellas funciones cuyas imágenes estén contenidas en {0,1} corresponden a los subconjuntos ordinarios o rígidos de U. De modo que un subconjunto ordinario o rígido, o simplemente un conjunto rígido, es un caso especial de un subconjunto borroso o, simplemente, de un conjunto borroso. En la literatura sobre conjuntos borrosos se acostumbra emplear dos notaciones: la letra A, llamada "etiqueta lingüística", y la letra griega, μ_A , que sería el análogo borroso a la función indicadora χ_A , por lo cual se conoce como función de pertenencia del conjunto borroso A. En este trabajo se hace la identificación $A = \mu_A$, i.e. se identifica cada conjunto borroso como su respectiva función de pertenencia, y, en concordancia, al valor $A(u) \in [0,1]$, definido para cada $u \in U$, se le llama el nivel de pertenencia de u en A.

Dado un concepto borroso, es claro que se pueden proponer distintas funciones A. esto significa que la elección de la función A es completamente subjetiva y dependiente del contexto. Este hecho se discute en el siguiente ejemplo.

¹ En su acepción en inglés se denominan conjuntos "crisp", es decir rígidos o crujientes, a los conjuntos definidos formalmente en la teoría de conjuntos, para distinguirlos de los conjuntos borrosos o vagos.

Ejemplo 1.1. Comparación de la función indicadora y de la función de pertenencia. Supongamos que se establece la clasificación de las estaturas de adultos varones mexicanos que se muestra en la tabla 1. Respecto a los datos mostrados en esta tabla, supongamos que las estaturas observadas en tres adultos mexicanos A, B y C son: 1.67 m, 1.69 m y 1.81 m, respectivamente. ¿Qué puede afirmarse acerca de este hecho empleando la función indicadora χ_A ?, y, ¿qué puede afirmarse si, en cambio, se utiliza la función de pertenencia μ_A ?; trataremos ahora de responder a estas preguntas básicas.

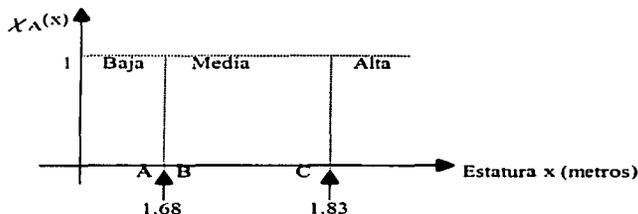
En la Tabla 2 se muestran los valores de la función indicadora para las tres personas mencionadas A, B y C; los cuales, a su vez, se observan en la Gráfica 1.

Estatura X (en metros)	Clasificación
$X < 1.68$	Baja
$1.68 \leq X \leq 1.83$	Media
$1.83 < X$	Alta

Tabla 1. Estatura (en metros) y clasificación arbitraria con tres etiquetas de hombres adultos mexicanos.

Persona	Estatura baja	Estatura media	Estatura alta
A	1	0	0
B	0	1	0
C	0	1	0

Tabla 2. Valores asignados por la función indicadora χ_A de acuerdo al Ejemplo 1, a tres personas A, B y C; cuyas estaturas son: 1.67 m, 1.69 m y 1.81 m, respectivamente.



Gráfica 1. Valores de la función indicadora χ_A para la estatura de tres adultos varones mexicanos del Ejemplo 1.

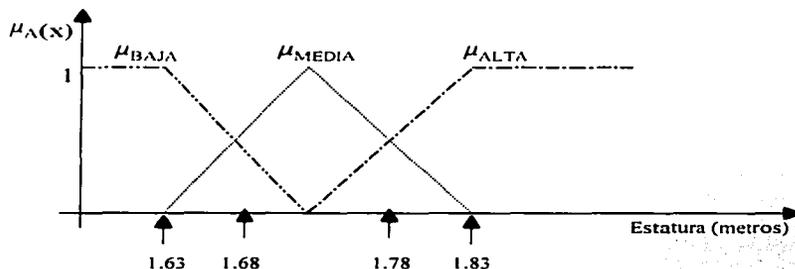
En la Tabla 3 se muestran las diferencias, en centímetros, encontradas en los valores de la estatura de los tres hombres adultos mexicanos del ejemplo 1. De acuerdo a la Tabla 3, B y C se clasifican como personas de estatura media mientras que A es clasificada como una persona de estatura baja, aunque la diferencia observada entre A y C es siete veces la correspondiente diferencia observada entre A y B.

Persona	A	B	C
A	0	2	14
B	2	0	12
C	14	12	0

Tabla 3. Diferencias, en centímetros, de las estaturas de A, B y C; de acuerdo al Ejemplo 1.

Esto confirma que las diferencias observadas entre los tres adultos A, B y C no son reflejadas apropiadamente al utilizar la función indicadora χ_A .

Ahora consideremos las funciones de pertenencia de tipo trapezoidal μ_A , donde $A = \{\text{"estatura baja", "estatura media", "estatura alta"}\}$, propuestas en la Gráfica 2.



Gráfica 2. Funciones de pertenencia propuestas para los datos del Ejemplo 1. se verifica que: $\mu_{BAJA}(x) = \mu_{MEDI A}(x)$, en $x = 1.68$ m; y, análogamente que: $\mu_{MEDI A}(x) = \mu_{ALTA}(x)$, en $x = 1.78$ m

Como puede observarse en la Gráfica 2, los valores de pertenencia $\mu_A(x)$ asociados a cada una de las tres personas A, B y C son los mostrados en la Tabla 4.

Persona	$\mu_{BAJA}(x)$	$\mu_{MEDI A}(x)$	$\mu_{ALTA}(x)$
A	0.6	0.4	0
B	0.4	0.6	0
C	0	0.2	0.8

Tabla 4. Valores de la función de pertenencia $\mu_A(x)$ asignados a las tres personas A, B y C del Ejemplo 1.

Si ahora se comparan los datos de la Tabla 2 y de la Tabla 4 es claro porqué el enfoque de los conjuntos borrosos nos permite expresar conceptos borrosos, como la estatura en este caso, de una manera más cercana a la empleada por el razonamiento humano que la utilizada en el enfoque de los conjuntos rígidos.

Aunque, como se mencionó previamente, en este trabajo se usa la notación $A: U \rightarrow [0,1]$, para representar a los conjuntos borrosos reales, y dicha representación no es única, se

pueden proponer diversas funciones de pertenencia A para representar a un conjunto borroso dado. Específicamente, en el Ejemplo 1 se pueden proponer alternativamente las funciones de pertenencia A_1 , de tipo trapezoidal, y A_2 , mezcla de trapezoidal y exponencial, que se definen enseguida:

$$A_1(x) = \begin{cases} 0 & \text{si } 0 \leq x < 1, \\ x-1 & \text{si } 1 \leq x < 2, \\ 1 & \text{si } 2 \leq x < 3, \\ 4-x & \text{si } 3 \leq x \leq 4, \\ 0 & \text{si } 4 < x. \end{cases} \quad (1.2)$$

$$A_2(x) = \begin{cases} \exp(x-3) & \text{si } 0 \leq x < 3, \\ 1 & \text{si } 3 \leq x < 5, \\ 1-(x-5)/2 & \text{si } 5 \leq x \leq 7, \\ 0 & \text{si } 7 < x. \end{cases} \quad (1.3)$$

Esta propuesta alternativa ilustra el hecho de que la selección de las funciones de pertenencia es un ejercicio de creatividad y de sensibilidad frente al contexto del concepto borroso que se quiere representar. Además, esta flexibilidad en la elección de las funciones de pertenencia resulta muy útil en caso de su aplicación al estudio de problemas reales, pues posibilita construir funciones de pertenencia razonables basadas, por ejemplo, en la información obtenida por medio de la realización de encuestas respecto al(los) concepto(s) específico(s) de interés. Por otra parte, es importante recalcar que el nivel de pertenencia $A(u)$ *no significa que conlleve ni la noción de probabilidad, ni la de verosimilitud, de que u posea el atributo o concepto borroso representado*. Más adelante se discute este punto nuevamente, pero antes se establecen las propiedades algebraicas básicas.

1.3 Operaciones básicas en los conjuntos borrosos

En esta sección definiremos las operaciones básicas definidas sobre los conjuntos borrosos de manera similar a las correspondientes a los conjuntos rígidos. Iniciaremos esta sección recordando a éstas últimas.

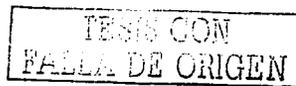
Sean A, B dos subconjuntos rígidos de un conjunto dado U con sus respectivas funciones indicadoras $\chi_A: U \rightarrow \{0,1\}$ y $\chi_B: U \rightarrow \{0,1\}$, y consideremos al conjunto potencia $\mathcal{P}(U)$ de todos los subconjuntos de U . Las operaciones de unión, intersección y complemento se definen en términos de las funciones indicadoras asociadas χ_A y χ_B como:

Unión: $\chi_{A \cup B}(x) = \text{máximo} \{ \chi_A(x), \chi_B(x) \} = \chi_A(x) \vee \chi_B(x)$

Intersección: $\chi_{A \cap B}(x) = \text{mínimo} \{ \chi_A(x), \chi_B(x) \} = \chi_A(x) \wedge \chi_B(x)$

Complemento: $\chi_{A^c}(x) = 1 - \chi_A(x)$.

Estas definiciones se generalizan de una manera natural a los subconjuntos borrosos de U de la siguiente manera, en donde $\mathcal{F}(U)$ es la familia de los subconjuntos borrosos de U .



Definición 1.2. Sean A, B conjuntos en $\mathcal{F}(U)$, con funciones de pertenencia $A: U \rightarrow [0,1]$ y $B: U \rightarrow [0,1]$. Entonces se definen las operaciones de unión, intersección y complemento de A y B como:

$$\text{Unión: } (A \vee B)(x) = \text{máximo } \{A(x), B(x)\} = A(x) \vee B(x) \quad (1.4)$$

$$\text{Intersección: } (A \wedge B)(x) = \text{mínimo } \{A(x), B(x)\} = A(x) \wedge B(x) \quad (1.5)$$

$$\text{Complemento: } A^c(x) = 1 - A(x). \quad (1.6)$$

La interpretación de la función $(A \vee B)(x)$ es que representa tanto al máximo como al supremo (sup) de las funciones $A(x)$ y $B(x)$, pues corresponde al valor mínimo que es mayor o igual tanto a $A(x)$ como a $B(x)$, i.e. es el supremo de ambas funciones. Esto es consistente con la notación empleada en la representación de retículas (lattices) que surgen al desarrollar las propiedades algebraicas de la TCB. Una interpretación análoga se tiene para la función $(A \wedge B)(x)$ como el mínimo y como el infimo (inf). Enseguida se propone un ejemplo para ilustrar el cálculo de estas operaciones básicas.

Ejemplo 1.2. Dadas las funciones de pertenencia A_1 y A_2 definidas en (1.2) y (1.3), respectivamente, determine su unión, su intersección y sus respectivos complementos.

$$(A_1 \vee A_2)(x) = \text{máximo } \{A_1(x), A_2(x)\} = \begin{cases} \exp(x-3) & \text{si } 0 \leq x \leq 1.158. \\ (x-1) & \text{si } 1.158 < x < 2. \\ 1 & \text{si } 2 \leq x < 5. \\ 1 - (x-5)/2 & \text{si } 5 \leq x \leq 7. \\ 0 & \text{si } 7 < x. \end{cases}$$

$$(A_1 \wedge A_2)(x) = \text{mínimo } \{A_1(x), A_2(x)\} = \begin{cases} 0 & \text{si } 0 \leq x < 1. \\ x-1 & \text{si } 1 \leq x < 1.158. \\ \exp(x-3) & \text{si } 1.158 \leq x < 3. \\ 4-x & \text{si } 3 \leq x < 4. \\ 0 & \text{si } 4 \leq x. \end{cases}$$

$$A_1^c(x) = 1 - A_1(x) = \begin{cases} 1 & \text{si } 0 \leq x < 1, \\ 2-x & \text{si } 1 \leq x < 2, \\ 0 & \text{si } 2 \leq x < 3, \\ x-3 & \text{si } 3 \leq x \leq 4, \\ 1 & \text{si } 4 < x. \end{cases}$$

$$A_2^c(x) = 1 - A_2(x) = \begin{cases} 1 - \exp(x-3) & \text{si } 0 \leq x < 3, \\ 0 & \text{si } 3 \leq x < 5, \\ (x-5)/2 & \text{si } 5 \leq x \leq 7, \\ 1 & \text{si } 7 < x. \end{cases}$$

1.4 Sobre los modelos matemáticos de la incertidumbre

Antes de plantear formalmente, en el siguiente capítulo, el problema de la clasificación de datos de tres formas (D3F), es conveniente caracterizar el problema de la clasificación de datos en general, desde el punto de vista de la base axiomática en la cual se sustentan dichos procesos. De acuerdo al tipo de incertidumbre implícito en cada tipo de procesos, éstos pueden ser: determinísticos, probabilísticos y borrosos.

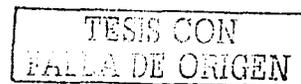
Un proceso es *determinístico* si sus resultados observados pueden ser predichos con toda certeza, en base a la replicación de las circunstancias que lo definen, por ejemplo, la segunda ley de Newton, o la ley de crecimiento exponencial. Estas leyes constituyen modelos determinísticos pues permiten estudiar las propiedades de los procesos determinísticos. En este tipo de procesos la única incertidumbre presente puede ser causada sólo como resultado de mediciones imprecisas, ya sea por parte del experimentador o por razones inherentes al uso de instrumentos de medición inapropiados, por ello se le conoce como *incertidumbre por imprecisión*. En este tipo de incertidumbre se incluye también la *imprecisión por ambigüedad*, la cual puede manifestarse en varias formas. Por ejemplo, consideremos un sistema de control en el cual sabemos que cierto parámetro P puede tomar valores dentro de un intervalo específico I . Entonces se tiene incertidumbre acerca del subconjunto de I cuyos valores puede tomar P .

En cambio, un proceso es *probabilístico* o aleatorio si sus correspondientes resultados u observaciones no pueden ser pronosticados con precisión, pues cambian aleatoriamente. La aleatoriedad se modela, generalmente, empleando la teoría de la probabilidad, es decir, se supone que los posibles resultados son observaciones de variables aleatorias, las cuales obedecen ciertas leyes de distribución cuyas propiedades matemáticas pueden ser establecidas teóricamente. Aunque estas leyes de distribución probabilística pueden no ser conocidas, cada variable aleatoria sigue una ley de distribución que es única, en contraste con el hecho de que se pueden asignar muchas funciones de pertenencia al mismo concepto borroso. Es decir, la probabilidad y los niveles de pertenencia corresponden a cosas completamente diferentes.

Por lo tanto, los modelos determinísticos y los modelos probabilísticos manejan distintos tipos y distintas cantidades de información de los respectivos procesos que representan, incluso en los casos en los que constituyan modelos del mismo proceso.

Se tienen además los procesos *borrosos* en los cuales los conceptos de estudio son de naturaleza borrosa o vaga: como educación, riqueza, salud, etc. En este caso los procesos modelados poseen otro tipo de incertidumbre no estadística, llamada borrosidad o vaguedad; distinta, esencialmente, a la incertidumbre por imprecisión y a la incertidumbre por aleatoriedad ya mencionadas. También se ha mencionado que la pertenencia a estos conceptos no es una cuestión de certeza, sino de grados o niveles, y cómo la propiedad de pertenencia se especifica matemáticamente por medio de los conjuntos borrosos.

Finalmente, se tiene el caso híbrido de los procesos *borrosos aleatorios*. Este tipo de procesos ocurre cuando se tienen muestras aleatorias y nos interesa estudiar conceptos o



atributos borrosos definidos sobre la población de la cual provienen las muestras. En este caso los conceptos borrosos de interés pueden ser modelados de manera rigurosa como conjuntos borrosos aleatorios.

En síntesis, cada tipo de incertidumbre tiene su propio modelo o representación matemática, y las correspondientes técnicas de cálculo asociadas. A final de cuentas, las diversas teorías matemáticas son como las distintas herramientas de una caja de herramientas, algunas son útiles para ciertas tareas o incluso puede ocurrir que, en ciertas condiciones, se desee emplear a todas en su conjunto. Por ejemplo, en áreas como el análisis de datos categóricos multivariados, la TCB puede proporcionar mejores resultados que las técnicas estadísticas convencionales, pues la TCB permite manejar categorías como "tensión mental", "nivel de depresión", "desequilibrio mental", etc. En la Tabla 5 se resumen, en general, los tipos de incertidumbre, sus características y las técnicas de modelaje matemático apropiadas.

Tipo de incertidumbre	Características	Modelo matemático
Imprecisión o ambigüedad	*Imprecisión instrumental *Ambigüedad en los valores que toman los objetos	Determinísticos Teoría moderna de control
Aleatoriedad	*Se desconocen los resultados en cada observación	Probabilísticos
Borrosidad	*Se desconocen con precisión cuánto de los atributos poseen los objetos.	Borrosos
Borrosa/Aleatoria	*Objetos borrosos definidos en muestras aleatorias	Borroso-aleatorios

Tabla 5. Características y técnicas de modelaje matemática empleadas para los distintos tipos de incertidumbre observados en el estudio de problemas de la vida real.

En el presente trabajo se propone un modelo de clasificación para datos de tres formas (D3F), los cuales definiremos posteriormente. El modelo propuesto emplea un método de clasificación basado en una función objetivo no lineal. Daremos en el siguiente capítulo un panorama de los diversos métodos de clasificación de datos, y plantharemos luego el problema al que nos enfrentamos al intentar clasificar este tipo de datos, y cómo resolverlo.

Capítulo 2. Un modelo multivariado de clasificación borrosa

TESIS CON
FALLA DE ORIGEN

2.1 Los métodos de clasificación de datos.

Empezaremos definiendo, en general, los distintos métodos de clasificación de datos. Posteriormente plantearemos el problema de la clasificación de datos en el contexto del reconocimiento de patrones y, finalmente, plantearemos el problema específico de clasificación borrosa de datos de tres formas (D3F) como un problema de optimización no lineal.

A grosso modo, se tienen tres tipos de métodos de clasificación de datos para cada clase axiomática de los modelos matemáticos subyacentes: jerárquicos, de teoría de gráficas y los que se basan en el uso de alguna función objetivo.

Los métodos de clasificación *jerárquicos* se originan en estudios de tipo taxonómico, especialmente en aquellos donde los datos de estudio poseen una estructura de tipo dendrítico, es decir, pueden acumularse o dividirse. Al hacerlo, se forman nuevas clases al reasignar la pertenencia de un dato específico en un momento dado, generando así una jerarquía de clases anidadas para cada clase inicial.

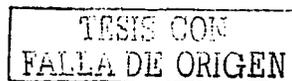
En los métodos de clasificación *basados en teoría de gráficas* se considera que el conjunto de datos X es un conjunto de nodos, con las ponderaciones en los extremos de cada uno construidas tomando en cuenta el grado de similitud entre cada par de nodos. Estos métodos son especialmente útiles en aquellos datos que están agrupados en "cadenas", o en estructuras de tipo lineal, o pseudolineal. Pero si los datos poseen estructuras mixtas, por ejemplo lineales y elípticas, estos métodos resultan inapropiados.

Finalmente, los métodos de clasificación basados en la utilización de alguna *función objetivo* permiten, generalmente, una formulación más precisa (aunque no necesariamente más válida) del criterio de clasificación. En este caso, un criterio o función objetivo mide la deseabilidad de los datos candidatos a ser incluidos en cada clase; y los valores extremos locales de esta función objetivo se definen como la clasificación óptima. Por ejemplo, si se toma como medida de similitud a la distancia euclidiana de los datos $X_i \in R^n$, y como medida de la bondad de la clasificación a la suma global intra-grupo de los errores al cuadrado (entre X_i y un valor central típico v_k), entonces la función objetivo es la suma de los errores cuadrados. Estos criterios son apropiados en aquellos casos donde los criterios basados en teoría de gráficas no lo son, Bezdek [1981].

Enseguida plantearemos el contexto donde surgen los D3F y los problemas que se presentan al intentar clasificarlos de una manera óptima.

2.2 Los datos de tres formas y su clasificación

En la investigación científica uno de los principales objetivos es determinar las relaciones empíricas que existen entre un conjunto dado de características. Con este fin se estudian las relaciones entre un conjunto de variables bajo ciertas condiciones específicas.



Las variables o atributos son las características que están siendo medidas y pueden ser cualquier cosa que pueda ser identificada o registrada de manera objetiva. Los objetos son los sujetos o unidades de estudio individuales que proporcionan los datos por medio de los cuales se evalúan las relaciones entre las variables. Las situaciones especifican las condiciones que son comunes a todos los datos recolectados en el estudio particular de interés; incluyen las posiciones en tiempo y espacio de los objetos de estudio, las variaciones en los métodos empleados para registrar las variables y los posibles tratamientos bajo los cuales se miden las variables o atributos. De esta manera una observación es el registro correspondiente a un atributo particular de un objeto específico obtenida bajo las situaciones establecidas.

Los datos considerados en el análisis de datos multidimensionales son, principalmente, de tipo matricial o datos de dos formas (D2F), los cuales consisten, en un primer caso, de objetos O y atributos A, por ejemplo, las mediciones de peso y estatura en niños. En este ejemplo sencillo, los niños son los objetos, y el peso y la estatura son los atributos. A estos datos los representamos abreviadamente como $O \times A$ (Objetos x Atributos). El segundo caso de los D2F consiste en una matriz de coeficientes de correlación o de covarianza entre atributos, la cual es utilizada en el análisis de componentes principales y en el análisis factorial, principalmente; este tipo de datos se representa abreviadamente como $A \times A$ (Atributos x Atributos). Análogamente, la matriz de las distancias entre objetos, que se representa como $O \times O$ (Objetos x Objetos), es utilizada en la realización del análisis de racimos (clusters), y en otros similares.

Los datos de 3 formas (D3F) consisten, generalmente, de objetos, atributos y situaciones, por ejemplo, las mediciones de peso y estatura de niños en diferentes edades. Representamos abreviadamente a los D3F como $O \times A \times S$, así que en este ejemplo O representa al conjunto de niños, A representa a sus atributos: peso y estatura, y S representa a sus distintas edades. Este tipo de D3F se muestra en la Figura 1.

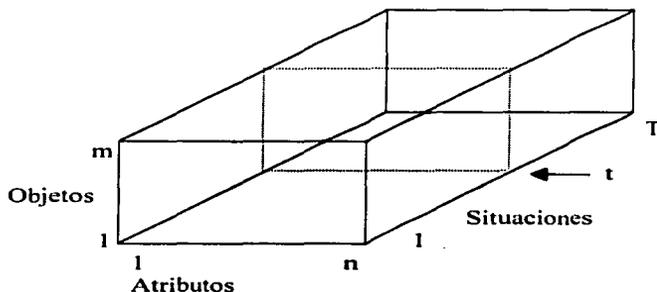


Figura 1. Representación de D3F del tipo $(O \times A \times S)$: m objetos, n atributos y T situaciones.

De manera similar al caso de los D2F, podemos mencionar algunos otros tipos de los D3F.

- (i) $A \times A \times S$: Son D3F que se representan por medio de una matriz de coeficientes de correlación entre los n atributos, para las S situaciones, como se muestra en la Figura 2.
- (ii) $A \times A \times O$: Consisten en D3F que se representan a través de una matriz de coeficientes de correlación entre los n atributos, respecto a los distintos objetos del estudio.
- (iii) $O \times O \times O$: Son D3F que representan la relación observada en triadas de los m objetos de estudio.

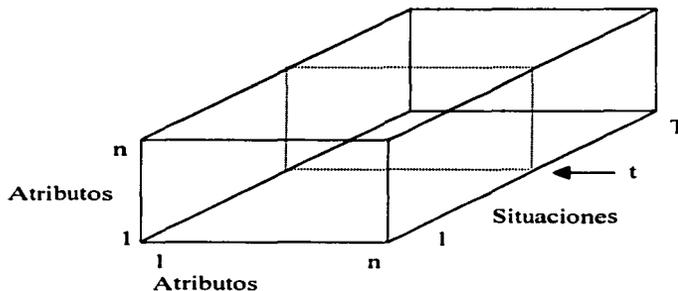


Figura 2. Representación de D3F del tipo $(A \times A \times S)$: Con n atributos y T situaciones.

Los D3F se emplearon inicialmente en el campo de la psicología cuantitativa, especialmente al estudiar los problemas del escalamiento multidimensional y del análisis factorial multimodal; por ejemplo, para D3F $O \times A \times S$ [Tucker, 1972], para D3F $O \times O \times O$ [Carroll-Arabie, 1983]. Sin embargo, en dichos modelos sólo se habían considerado variables rígidas convencionales. La aportación al respecto del presente trabajo abre una amplia gama de perspectivas de aplicación pues permite mejorar las investigaciones en diseños experimentales longitudinales, es decir realizados a lo largo de distintas etapas temporales.

En este trabajo se propone un modelo para clasificar D3F, particularmente del tipo $O \times A \times S$, cuyo objetivo es revelar la estructura latente de este tipo de datos a través de todos los puntos del tiempo (situaciones) por medio de la construcción de *clases* que consideren no sólo la similitud entre un par de objetos en cada tiempo, sino también la similitud en el patrón de cambio de las observaciones con respecto al tiempo. Sin embargo, en D3F la clasificación de objetos presenta el problema de que una posible clasificación de objetos O no es, generalmente, invariable con respecto a todas las situaciones S , debido a que los valores de los correspondientes atributos A cambian con cada nueva situación. Es decir que en D3F las clases óptimas para cada punto temporal no son siempre los mismos respecto a un determinado criterio, y no es recomendable aceptar sólo el resultado para un punto temporal específico.

Debemos entonces encontrar una solución que sea consistente con todos los criterios y que satisfaga a todos ellos tanto como sea posible. Por lo tanto, el problema de clasificación de D3F puede considerarse como un problema de *optimización de criterios múltiples* y, consecuentemente, lo abordaremos desde esta perspectiva.

2.3 Conceptos básicos del reconocimiento de patrones

Estableceremos ahora la conexión del problema general de clasificación de datos en el contexto del reconocimiento de patrones, con el fin de ubicar apropiadamente el problema particular de la clasificación para D3F en dicho contexto.

Se realiza un proceso de *reconocimiento* cuando se determina que un objeto de una población P pertenece a una subpoblación S. La *identificación* es el reconocimiento de un objeto individual como una clase única de un solo elemento. La *clasificación* consiste en el proceso de agrupación de objetos en clases (subpoblaciones) de acuerdo a su parecido o similitud.

El área de conocimiento denominado *reconocimiento de patrones* incluye tanto la clasificación como el reconocimiento y pertenece al campo más amplio conocido como máquinas inteligentes, es decir, el estudio y desarrollo de técnicas para lograr que las máquinas aprendan y razonen para tomar decisiones, tal como lo hacen los humanos.

Dada una población P, los atributos de un objeto son detectados o medidos para obtener un *vector patrón*, el cual es transformado en un conjunto reducido de características llamado *vector característico*. Un *sistema de reconocimiento de patrones* consiste en un sistema al que se introduce un vector característico como entrada, sobre el que opera el sistema en la fase llamada *modo operacional*, para producir una salida que es el *identificador* único asociado con la clase a la cual pertenece el objeto.

Sin embargo, antes de que el sistema pueda tomar la decisión sobre cuál identificador asociar al vector característico de entrada, el sistema debe haber aprendido, en la fase llamada *modo de entrenamiento*, las clases de vectores característicos. Esto se logra por medio de un procedimiento de partición del conjunto de vectores característicos denominado *clasificación*, el cual supone el concepto de *entrenamiento o aprendizaje*.

Las categorías de aprendizaje de las máquinas son dos: *entrenamiento supervisado*, en el cual el sistema aprende a mapear el vector característico de entrada hacia el identificador de clase deseado; y *entrenamiento no supervisado o auto-organización*, en el cual el sistema realiza los modos de entrenamiento y operacional sobre cualquier vector característico de entrada. En el modo auto-organizado el proceso de agrupar vectores característicos de entrada en clases (racimos) se llama clasificación (agrupamiento), en inglés "clustering". Existe una amplia bibliografía sobre los temas de reconocimiento de patrones y la clasificación de datos, empleando específicamente técnicas recientes como las basadas en redes neuronales y lógica borrosa [Looney, 1997].

El principal objetivo de la clasificación no supervisada de un conjunto de objetos es detectar subgrupos que existen de manera natural, basados en la posible similitud o disimilitud observada en un par de objetos. Específicamente, al enfocar el problema de clasificación de D3F como un problema de optimización de criterios múltiples, esperamos que el presente trabajo se pueda considerar como un paso inicial hacia la construcción de un sistema de reconocimiento de patrones para D3F.

TESIS CON
FALLA DE ORIGEN

2.4 El problema de optimización de criterios múltiples

El problema de optimización de un solo criterio (POSC), representado como (Ω, J) , se caracteriza al plantear como objetivo encontrar la solución x^* para la cual se verifica que

$$J(x^*) = \text{mínimo } \{J(x) \mid x \in \Omega\} \quad (2.1)$$

donde $J: \Omega \rightarrow \mathbf{R}$ es la (única) función criterio y $\Omega \subset \mathbf{R}^n$ es el conjunto de soluciones factibles. Por otra parte, en el problema de optimización de criterios múltiples (POCM), representado como $(\Omega, J_1, J_2, \dots, J_T)$, el objetivo es determinar la solución $x^* \in \Omega$ para la cual

$$J_t(x^*) = \text{"mínimo"} \{J_t(x) \mid x \in \Omega\} \quad (2.2)$$

donde $J = (J_1, \dots, J_T): \Omega \rightarrow \mathbf{R}^T$ es el vector de las T funciones criterio definidas sobre el conjunto de soluciones factibles Ω . Podemos suponer que las funciones $J_t(x)$ son todas acotadas inferiormente, específicamente supondremos entonces que $J_t(x) > 0$, para todo t . Sin embargo, como J es un vector de funciones, el concepto de "mínimo" en (2.2) no tiene un significado único. En la caracterización de la solución del POCM se emplean varios conceptos como: solución dominante, solución dominada y solución eficiente o de Pareto, que ahora definimos.

Definición 1. Solución dominante

Se dice que $x_0 \in \Omega$ es una solución dominante si se verifica que:

$$J_t(x_0) = \text{mínimo } \{J_t(x) \mid x \in \Omega\} \quad \forall t = 1, 2, \dots, T.$$

Es decir, una solución es dominante si es óptima para cada una de las T funciones objetivo J_t simultáneamente. No obstante, como en la mayoría de los casos no existe una solución dominante, se deben buscar soluciones alternativas. Tales soluciones se llaman soluciones eficientes o de Pareto, y su definición está asociada con el concepto de soluciones dominadas que definimos enseguida.

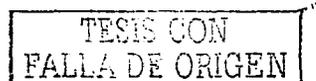
Definición 2. Solución dominada

Dadas dos soluciones x_1 y x_2 , si $J_t(x_1) \leq J_t(x_2)$ para todo $t \in \{1, 2, \dots, T\}$, y si además se cumple que $J_s(x_1) < J_s(x_2)$ para al menos una $s \in \{1, 2, \dots, T\}$, entonces decimos que la solución x_1 domina a la solución x_2 , o, equivalentemente, que x_2 es dominada por x_1 . En este caso escribimos $J(x_1) < J(x_2)$.

La diferencia entre una solución dominante y una solución dominada es que la primera es un concepto absoluto y la segunda es un concepto relativo definido para parejas de soluciones.

Definición 3. Solución eficiente

Una solución $x^* \in \Omega$ es una solución eficiente (o de Pareto) si no existe otra solución $x \in \Omega$ que la domine. Es decir, no existe $x \in \Omega$ tal que $J(x) < J(x^*)$.



Definición 4. Solución eficiente local

Decimos que la solución $x^* \in \Omega$ es una solución eficiente (o de Pareto) local si existe una vecindad abierta $N(x^*)$ de x^* tal que x^* es eficiente en $\Omega \cap N(x^*)$; es decir, no existe solución $x \in \Omega \cap N(x^*)$ que domine a x^* .

Es posible obtener soluciones eficientes (SE) al transformar el POCM en ciertos POSC. Al respecto se han desarrollado diversos métodos de transformación, por ejemplo el método de sumas ponderadas [Da Cunha-Polak, 1967], el método del punto de referencia ideal [Yu, 1973], y el método minimax ponderado [Bowman, 1976]. En este trabajo consideraremos el método de sumas ponderadas pues facilita la conexión del POCM al contexto de los D3F.

Definición 5. Transformación de un POCM en un POSC con el método de sumas ponderadas

Sea $\Omega_T \subset \mathbb{R}^T$ el conjunto de vectores $\omega \in \mathbb{R}^T$ con componentes $\omega_t > 0$ para todo $t=1,2,\dots,T$ y con $\omega_1 + \omega_2 + \dots + \omega_T = 1$. Dado un POCM $(\Omega, J_1, \dots, J_T)$, para cada $\omega \in \Omega_T$ se define la función $P_{\omega}(x) = \sum_1 \omega_t J_t(x)$ para $x \in \Omega$. En este caso se dice que (Ω, P_{ω}) es el POSC asociado al POCM y a $\omega \in \Omega_T$. Nótese que $P_{\omega}(x) = \omega \cdot J(x)$, el producto escalar de los vectores $\omega = (\omega_1, \omega_2, \dots, \omega_T)$ y $J(x) = (J_1(x), \dots, J_T(x))$.

Al emplear esta transformación surge el problema de cómo garantizar que al resolver el POSC se obtenga también una solución del POCM original, y viceversa. Veamos enseguida el resultado teórico que resuelve este problema [Da Cunha-Polak, 1967], [Shapiro, 1979].

Teorema 1. (a) Sea $x^* \in \Omega$ una solución óptima del POSC (Ω, P_{ω}) con vector de ponderaciones $\omega \in \Omega_T$. Entonces x^* es una solución eficiente del POCM $(\Omega, J_1, J_2, \dots, J_T)$.
 (b) Recíprocamente, si cada J_t , $t=1,2,\dots,T$, es una función convexa, y si x^* es solución eficiente del POCM $(\Omega, J_1, J_2, \dots, J_T)$, entonces existe $\omega \in \Omega_T$ tal que x^* es una solución óptima del POSC (Ω, P_{ω}) .

Demostración.

(a) Supongamos que x^* no es una SE del POCM. De acuerdo con la definición 3 esta suposición significa que existe una solución $x \in \Omega$ tal que $J(x) < J(x^*)$.

Por lo tanto, $P_{\omega}(x) = \omega \cdot J(x) < \omega \cdot J(x^*) = P_{\omega}(x^*)$, y se contradice así que x^* es una solución óptima del POSC (Ω, P_{ω}) .

(b) Sea x^* una solución eficiente del problema $(\Omega, J_1, J_2, \dots, J_T)$, se definen los conjuntos:

$G = \{g \in \mathbb{R}^T \mid g = (g_t), \text{ con } g_t > 0, t=1,2, \dots, T\}$,

$H = \{h \in \mathbb{R}^T \mid h = (h_t), \text{ con } h_t \leq J_t(x^*) - J_t(x) \ \forall x \in \Omega, t=1,2,\dots,T\}$

Se comprueba entonces que G y H son conjuntos ajenos y que ambos son convexos. luego, por el teorema de separación de conjuntos convexos [Rockafeller, 1970], existe $\omega \in \mathbb{R}^{T(+)}$ tal que: $\omega \cdot h \leq 0 < \omega \cdot g$ para todo $g \in G, h \in H$.

En particular, si tomamos $h_t = J_t(x^*) - J_t(x)$, entonces se verifica que: $\sum_1 \omega_t h_t = \sum_1 \omega_t (J_t(x^*) - J_t(x)) \leq 0$, y en consecuencia, $\sum_1 \omega_t J_t(x^*) \leq \sum_1 \omega_t J_t(x)$, para todo $x \in \Omega$.

Finalmente, denotando por $\|\omega\|$ la norma de ω , y definiendo $\hat{\omega} := \omega / \|\omega\|$, vemos que $\hat{\omega} \in \Omega_T$ satisface la conclusión en (b) ■

2.5 El problema de clasificación para D3F

Representaremos los D3F que se tienen al observar en m objetos los valores de n atributos respecto a T situaciones como:

$$X^{(t)}_{ij}, \text{ donde } 1 \leq i \leq m, 1 \leq j \leq n, 1 \leq t \leq T \quad (2.3)$$

El objetivo de este trabajo es construir un modelo para clasificar D3F como (2.3) respecto a K clases borrosas C_k , $1 \leq k \leq K$. Estas clases borrosas son subconjuntos del conjunto de objetos $O = \{o_i \in \mathbb{R}^n \mid 1 \leq i \leq m\}$, en el sentido borroso que se define en el capítulo 1.

Se representa el grado de pertenencia del objeto o_i a la clase C_k , como μ_{ki} , el elemento (ki) -ésimo de la matriz borrosa $(K \times m)$ dada por:

$$U = (\mu_{ki}) \text{ con } \mu_{ki} \geq 0 \text{ y } \sum_k \mu_{ki} = 1 \forall i, \text{ con } 1 \leq k \leq K, 1 \leq i \leq m. \quad (2.4)$$

Para cada clase, se define su respectivo "centro de masa", por analogía con el concepto físico, comúnmente llamado centroide, asociado al tiempo t como:

$$v^{(t)}_k = (v^{(t)}_{kj}) \text{ para } 1 \leq k \leq K, 1 \leq j \leq n, 1 \leq t \leq T \quad (2.5)$$

donde $v^{(t)}_{kj}$ es la j -ésima componente del centroide $v^{(t)}_k$ sobre \mathbb{R}^n .

Matemáticamente el problema principal para lograr una clasificación de D3F consiste en determinar la matriz de pertenencias $U = (\mu_{ki})$ y sus correspondientes centroides $v^{(t)}_k$.

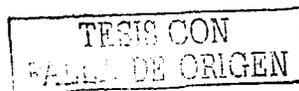
Finalmente, siguiendo a Bezdek [Bezdek, 1981], la medida de la bondad de clasificación al tiempo t , considerando un nivel de borrosidad α en el modelo de clasificación, está dada por la suma generalizada de la dispersión intra-clase:

$$J^{(t)}(U, v^{(t)}) = \sum_k \sum_i (\mu_{ki})^\alpha d^2(o_i, v^{(t)}_k), \quad (2.6)$$

donde $d(o_i, v^{(t)}_k) = [\sum_j (X^{(t)}_{ij} - v^{(t)}_{kj})^2]^{(1/2)}$ es la distancia euclidiana de o_i a $v^{(t)}_k$.

Es claro en (2.6) que $J^{(t)}$ es una función de cuadrados mínimos definida como una medida de la suma de las distancias euclidianas al cuadrado ponderadas con los valores de pertenencia $(\mu_{ki})^\alpha$. Respecto al parámetro α , es utilizado para controlar la cantidad de borrosidad del proceso de clasificación, y su rango de valores es $[1, \infty)$. Cuando, en particular $\alpha=1$, y en (2.6) no se considera la evolución temporal del proceso, entonces (2.6) se reduce al algoritmo de clasificación para datos rígidos o convencionales; si, en el otro extremo, $\alpha \rightarrow \infty$, y se mantiene constante el tiempo t , entonces $J^{(t)} \rightarrow 0$ [Bezdek, 1981]. Por lo tanto, α controla la extensión de la pertenencia compartida entre las diversas clases definidas en el proceso. No se ha reportado hasta ahora un valor óptimo para α , aunque la mayoría de los trabajos emplean $\alpha \in [1.25, 2]$, Ross [1995]. Más adelante se examina un ejemplo empleando el valor $\alpha = 2$, el cual ha sido empleado en la mayoría de los trabajos publicados a la fecha [Dunn, 1974], [Bezdek, 1981], con resultados aceptables para D2F.

Debe observarse también que las clases, que son subconjuntos borrosos de O , se determinan unívocamente a lo largo del tiempo t , $t = 1, 2, \dots, T$; sin embargo, de acuerdo con (2.5), los centroides sí cambian con respecto al tiempo. Lo anterior significa que el objetivo del procedimiento de clasificación borrosa representado en (2.6) es construir clases



comunes e independientes del tiempo, es decir, se busca obtener una matriz borrosa $U=(\mu_{ki})$, donde los elementos μ_{ki} no dependan del tiempo t . Supongamos que no es así, es decir, supongamos que en el problema convencional de clasificación de D3F las clases construidas C_k fueran dependientes de t . En este caso se tendría una clasificación convencional rígida para los D2F, obtenidos al mantener constante t en cada situación. Por lo tanto, en este caso, no es necesario considerar el POCM dado representado en (2.6), ni tampoco resolverlo, pues ya ha sido resuelto previamente [Bezdek, 1981].

Consideremos ahora, específicamente, el problema de encontrar una solución $(U, \nu)^*$ para el POCM definido por los criterios en (2.6), es decir,

$$J((U, \nu)^*) = \text{mínimo } \{J^{(t)}(U, \nu^{(t)}) \mid t=1, 2, \dots, T\}. \quad (2.7)$$

De acuerdo con la definición 1, $(U, \nu)^*$ sería la solución dominante óptima; sin embargo tal solución no existe o es muy difícil garantizar que lo sea. Para aclarar este punto consideremos el concepto de partición.

Definición 6. Sean $X=\{X_1, \dots, X_m\}$ cualquier conjunto finito, $V_{K \times m}$ el conjunto de matrices reales ($K \times m$), K un entero con $2 \leq K < m$. Entonces el espacio de K -particiones rígidas para X es el conjunto

$$M_K = \{U \in V_{K \times m} \mid \chi_{ki} \in \{0, 1\}, \forall i, \forall k; \sum_k \chi_{ki} = 1 \forall k; 0 < \sum_i \chi_{ki} < n \forall i\} \quad (2.8)$$

donde $\chi_{ki} = \chi_k(X_i)$ es el valor de la función indicadora de la clase k evaluada en X_i .

La generalización de este concepto al caso de K -particiones borrosas es [Ruspini, 1970]:

Definición 7. Sean $X=\{X_1, \dots, X_m\}$ cualquier conjunto finito, $V_{K \times m}$ el conjunto de matrices reales ($K \times m$), K un entero con $2 \leq K < m$. Entonces el espacio de K -particiones borrosas para X es el conjunto:

$$M_{bK} = \{U \in V_{K \times m} \mid \mu_{ki} \in [0, 1] \forall i, \forall k; \sum_k \mu_{ki} = 1 \forall k; 0 < \sum_i \mu_{ki} < n\} \quad (2.9)$$

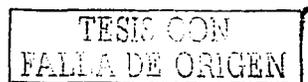
donde $\mu_{ki} = \mu_k(X_i)$ es el valor de la función de pertenencia de X_i a la clase k .

De (2.8) podemos inferir que M_K , el espacio de K -particiones rígidas, es finito pero su cardinalidad $|M_K|$ es muy grande para casi todos los valores de K y m , excepto para los casos triviales. De hecho, se puede demostrar que el número de maneras distintas de particionar X , formado por m objetos, en K subconjuntos rígidos no vacíos es:

$$|M_K| = (1/K!)^{-1} \sum_x \binom{K}{x} (-1)^{(K-x)} (x)^m, \quad \kappa = 1, 2, \dots, K. \quad (2.10)$$

con la condición obvia $2 \leq K < m$, donde $\binom{K}{x}$ es el coeficiente binomial [Ross, 1995].

Por ejemplo, si $m=20$ y $K=2$, entonces $|M_2| = 524287$; análogamente, para $m=20$ y $K=3$, se obtiene $|M_3| = 580606446$. Como la cardinalidad $|M_K|$ es muy grande, aún para valores reducidos de m y de K , el procedimiento de solución, al tener que recorrer un gran espacio



de soluciones factibles, se convierte en un problema de optimización combinatoria. En cambio, al definir (2.6) haciendo que $J^{(0)}(\bullet)$ dependa explícitamente de (μ_{ki}) se logra que $J^{(0)}(\bullet)$ sea diferenciable sobre M_{bk} , pero no sobre M_K ; lo cual permite resolver (2.7).

Haremos ver ahora cómo obtener SEs para resolver el problema de la clasificación borrosa con criterios múltiples para D3F. La demostración que presentaremos se logra al generalizar el algoritmo de clasificación borroso para D2F [Bezdek, 1981]. Es conveniente, por lo tanto, presentar primero este resultado fundamental para el presente trabajo.

Definición 8. Representaremos los D2F que se tienen al observar o medir en m objetos los valores de n atributos como:

$$X_{ij} \text{ para } 1 \leq i \leq m, 1 \leq j \leq n. \quad (2.11)$$

El objetivo es construir un modelo para clasificar D2F como (2.9) respecto a K clases borrosas C_k , $1 \leq k \leq K$. Estas K clases borrosas son subconjuntos del conjunto de objetos $O = \{o_i \mid 1 \leq i \leq m\}$. Se representa el grado de pertenencia del objeto o_i a la clase C_k como μ_{ki} , el elemento (ki) -ésimo de la matriz borrosa $(K \times m)$:

$$U = (\mu_{ki}), \text{ en la cual } \mu_{ki} \geq 0, \sum_k \mu_{ki} = 1, 1 \leq k \leq K, 1 \leq i \leq m. \quad (2.12)$$

Para cada clase k , su elemento representativo, denominado centroide, se define como:

$$v_k := (v_{k1}, \dots, v_{kn}) \in R^{nK} \text{ con } v_k = (v_{k1}, \dots, v_{kn}) \in R^n \quad \forall k. \quad (2.13)$$

Finalmente, para las funciones $J_m: M_{bK} \times R^{nK} \rightarrow R^+$, la medida de la bondad de clasificación se define en forma similar a (2.6), es decir

$$J_\alpha(U, v) = \sum_k \sum_i (\mu_{ki})^\alpha d^2(o_i, v_k), \quad (2.14)$$

donde $U \in M_{bK}$ y $v_k = (v_{k1}, \dots, v_{kn}) \in R^n \quad \forall k$,

Debido a que cada término de $J_\alpha(U, v)$ es proporcional a $d^2(x_i, v_k)$, entonces J_α es una función objetivo que define un criterio de clasificación de error cuadrático. Por lo tanto, las soluciones $(U, v)^*$ del problema de programación matemática, para las cuales:

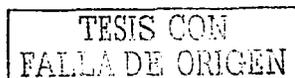
$$J((U, v)^*) = \text{Mínimo} \{J_\alpha(U, v) \mid U \text{ satisface (2.12)}\} \quad (2.15)$$

son puntos estacionarios de cuadrados mínimos de $J_\alpha(U, v)$. Para cada valor $\alpha \in [1, \infty)$ se obtiene entonces una familia infinita de algoritmos de clasificación borrosa al obtener las condiciones necesarias para resolver (2.15). Esto se logra con el siguiente resultado [Bezdek, 1981]

Teorema 2. Sean $\|\bullet\|$ cualquier norma de producto interno inducida sobre R^n y $\alpha \in [1, \infty)$ un valor fijo, y supongamos que el conjunto de datos X tiene al menos K puntos distintos, $2 \leq K < m$. Se definen, además, para $i=1, 2, \dots, m$; los conjuntos:

$$C_i = \{k \mid d(o_i, v_k) = \|o_i - v_k\| = 0, 1 \leq k \leq K\}$$

$$C_i^c = \{1, 2, \dots, K\} - C_i, \text{ el complemento borroso de } C_i.$$



Entonces $(U, \nu)^*$ es una solución óptima de (2.15) siempre y cuando:

$$\nu_k = [\sum_i (\mu_{ki})^\alpha \mathbf{X}_i] / [\sum_i (\mu_{ki})^\alpha] \quad \forall k=1,2,\dots,K \quad (2.16)$$

$$\text{con } C_i = \emptyset \text{ si } \mu_{ki} = \left[\sum_{\delta=1}^K (d_{ki}/d_{\delta i})^{2/(k-1)} \right]^{-1}, \text{ donde } d_{ki}=d(o_i, \mu_k) \quad (2.17a)$$

$$\text{y } C_i \neq \emptyset \text{ si } \mu_{ki} = 0 \quad \forall k \in C_i^c \text{ y } \sum_{k \in C_i} \mu_{ki} = 1 \quad (2.17b)$$

Aunque no demostraremos aquí este teorema, describiremos *a grosso modo* cómo se realiza la demostración porque en este trabajo usaremos argumentos semejantes. La ecuación (2.16) se deduce directamente por medio de la aplicación de técnicas de optimización no restringida a $J_\alpha(U, \nu)$, considerando como variables a $\{\nu_k\}$ y manteniendo fijas las variables (μ_{ki}) . Las ecuaciones (2.17) se obtienen al mantener fijas las $\{\nu_k\}$ y aplicando entonces la técnica de los multiplicadores de Lagrange a la función objetivo resultante de las variables (μ_{ki}) . Estos argumentos constituyen entonces la pauta para encontrar, enseguida, SEs al POCM dado en (2.7).

Antes de formular el teorema para clasificar D3F, debemos hacer dos consideraciones importantes que son consecuencias del Teorema 2.

En primer lugar, las ecuaciones (2.16) y (2.17) se cumplen para cualquier norma inducida por un producto interno. En particular, cualquier matriz positiva definida A , sobre V_{int} induce una norma por medio del producto interno ponderado:

$$\langle \mathbf{x}, \mathbf{y} \rangle_A = \mathbf{x}^t \mathbf{A} \mathbf{y} = \sum_i \sum_j x_i a_{ij} y_j, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \quad (2.18)$$

En este tipo especial de normas, la función $J_\alpha(U, \nu)$ en (2.14) puede expresarse como:

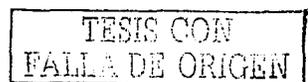
$$J_\alpha(U, \nu, A) = \sum_k \sum_i (\mu_{ki})^\alpha \|\mathbf{x}_i - \nu_k\|_A^2. \quad (2.19)$$

En donde $d^2(\mathbf{x}_i, \nu_k) = \|\mathbf{x}_i - \nu_k\|_A^2 = \langle \mathbf{x}_i - \nu_k, \mathbf{x}_i - \nu_k \rangle_A = (\mathbf{x}_i - \nu_k)^t \mathbf{A} (\mathbf{x}_i - \nu_k)$. En esta expresión queda explícita la dependencia de J_α respecto a la matriz A que define a la norma en \mathbb{R}^n a través de (2.18). Esto significa que, en general, la matriz A es una variable de tipo algorítmico para los diversos métodos de clasificación basados en una función objetivo como la definida en el Teorema 2.

En segundo lugar, el algoritmo de clasificación borrosa para datos de dos formas, que se establece como consecuencia del Teorema 2, es un procedimiento iterativo que consta de cuatro fases, el cual planteamos enseguida.

Algoritmo de clasificación de K-medias borrosas

Fase 1. Se determina el número de clases K , con la condición obvia: $2 \leq K < m$; se elige cualquier norma sobre \mathbb{R}^n inducida por un producto interno; se fija el valor del parámetro α , tal que $1 \leq \alpha < \infty$. Se propone una matriz borrosa inicial $U^{(0)}$. Luego, para el paso r , $r = 0, 1, \dots$



Fase 2. Se calculan los centroides de las K clases borrosas $\{v_k^{(t)}\}$ empleando (2.16) así como los elementos $(\mu^{(t)})_{ki}$ de la matriz $U^{(t)}$.

Fase 3. Se actualiza $U^{(t)}$, utilizando (2.17) y $\{v_k^{(t)}\}$

Fase 4. Se comparan las matrices $U^{(t)}$ y $U^{(t+1)}$. Hacer alto si, dado un valor $\epsilon_t > 0$ previamente establecido: $\|U^{(t)} - U^{(t+1)}\| \leq \epsilon_t$; en caso contrario, regresar a la fase 2. ■

Proponemos ahora el resultado que permite, de manera análoga al caso del Teorema 2, obtener SEs del POCM (2.6), este resultado expresado en forma de teorema es la principal aportación de este trabajo.

Supondremos nuevamente que Ω es el conjunto de soluciones factibles de la función objetivo $J^{(t)}(U, v^{(t)})$ en (2.6), es decir que:

$$\Omega = \{(U, v) | U = (\mu_{ki}), v = (v_k^{(t)}), 1 \leq k \leq K, 1 \leq i \leq m, 1 \leq t \leq T\}.$$

De acuerdo con la Definición 5, $J(U, v)$ es la función criterio de las ponderaciones dada por:

$$J(U, v) = \sum_i \omega^{(i)} J^{(i)}(U, v_k^{(i)}) \text{ con } \omega^{(i)} > 0 \quad (2.20)$$

Por el Teorema 1 el problema de obtener una SE del POCM $(\Omega, J^{(1)}, \dots, J^{(T)})$ se transforma en un problema de optimización no lineal. Encontraremos ahora las condiciones necesarias para obtener, de acuerdo a la Definición 4, las SEs locales de $(\Omega, J^{(1)}, \dots, J^{(T)})$, considerando que $J(U, v)$ en (2.20) es una función criterio continua de variables múltiples.

Teorema 3. Supongamos que se dispone de la matriz de D3F definida en (2.3): $X^{(t)} = (X^{(t)})_{ij}$, con $1 \leq i \leq m, 1 \leq j \leq n, 1 \leq t \leq T$; y supongamos que los m objetos se clasifican en K clases borrosas: C_1, \dots, C_K . Sea $U = (\mu_{ki})$ la matriz borrosa de pertenencias, es decir, μ_{ki} es el nivel de pertenencia del objeto o_i a la clase C_k , con $v_k^{(t)}$ siendo el centroide de la clase C_k al tiempo t . Sea (U^*, v^*) una SE del POCM $(\Omega, J^{(1)}, \dots, J^{(T)})$, es decir

$$J(U^*, v^*) = \text{mínimo}_{(U, v)} \{J(U, v)\}. \quad (2.21)$$

Se definen ahora los siguientes conjuntos:

$$C_i = \{k=1, \dots, K | \sum_t d^2(X^{(t)})_{ij}, v_k^{(t)}\} = \sum_t \sum_j (X^{(t)})_{ij} - v_k^{(t)}\}^2 = 0\}$$

$$C_i^c = \{k=1, \dots, K\} - C_i, \text{ el complemento borroso de } C_i.$$

Entonces: (a) Dados los valores de pertenencia de $U = (\mu_{ki})$, se tiene que

$$v_{kj}^{(t)} = [\sum_i (\mu_{ki})^\alpha X_{ij}^{(t)}] / [\sum_i (\mu_{ki})^\alpha], \text{ para } 1 \leq k \leq K, 1 \leq j \leq n, 1 \leq t \leq T \quad (2.22)$$

(b) Dados los valores de los centroides $v_{kj}^{(t)}$, se tiene que

$$\mu_{ki} = \begin{cases} \left[\sum_{q=1}^K \frac{[\sum_t \omega^{(t)} d^2(X^{(t)})_{ij}, v_k^{(t)}]}{[\sum_t \omega^{(t)} d^2(X^{(t)})_{ij}, v_q^{(t)}]} \right]^{1/(\alpha-1)} - 1 & \text{si } C_i = \emptyset, \quad (2.23a) \\ 0 & \forall k \in C_i^c, \sum_k (\mu_{ki}) = 1, k \in C_i \quad \text{si } C_i \neq \emptyset. \quad (2.23b) \end{cases}$$

Análisis del teorema.

La evidente similitud de los Teoremas 2 y 3 nos permite seguir el procedimiento empleado en la demostración del Teorema 2 para demostrar el Teorema 3. Las ecuaciones (2.22) y (2.16) son idénticas; de hecho se obtienen por medio de procedimientos de optimización sin restricciones aplicados a las funciones objetivo (criterios) J_α y $J^{(t)}$, respectivamente. Respecto a la dependencia en el tiempo de $J^{(t)}$, y por lo tanto de $J(U, \nu)$ en (2.21), el objetivo de este teorema es obtener clases comunes válidas para todos los tiempos de tal manera que, consecuentemente, los correspondientes centros de cada clase asociados a la SE caracterizada por los valores de pertenencia μ_{ki} sean también independientes del tiempo.

En la ecuación (2.23b), se observa que $\mu_{ki}=0$ si existe k tal que $d(X^{(t)}, \nu_k^{(t)})=0$. El objeto o_i que satisface esta condición debe tener nivel de pertenencia cero en cualquier otra clase, y por esta razón se le conoce como *singularidad*.

La función objetivo no lineal $J(U, \nu)$ en (2.20) se considera una función continua de las variables múltiples (U, ν) para garantizar que las ecuaciones (2.22), (2.23a) y (2.23b) permitan alcanzar una SE en (2.21).

También es conveniente enfatizar que la calidad del procedimiento de clasificación depende de la elección apropiada de los parámetros α y ω . En primer lugar, respecto al parámetro real $\alpha \in [1, \infty)$, que mide el nivel de borrosidad del modelo, utilizaremos el valor $\alpha=2$, que permite lograr un criterio de clasificación aceptable, como se ha reportado en trabajos similares [Dunn, 1974], [Ross, 1995]. En segundo lugar, respecto a los factores de ponderación ω , en el estudio de procesos en los cuales no se disponga de información sobre cuáles de las situaciones son más relevantes para determinar la influencia de los n atributos sobre los m objetos, se deben asignar pesos equitativos, y las excepciones a esta equidad en los factores de ponderación ω sólo se justifican cuando se disponga de este tipo de información.

Demostración del Teorema 3.

Primero calcularemos (2.22) suponiendo que la matriz de pertenencias borrosas $U=(\mu_{ki})$ es fija, con los valores μ_{ki} independientes del tiempo. Entonces:

$$\begin{aligned} \text{Mínimo}_{\{U, \nu\}} J(U, \nu) &= \text{Mínimo}_{\{U, \nu\}} \{ \sum_{i=1}^T \omega^{(t)} J^{(t)}(U, \nu_k^{(t)}) | \nu_k \in \mathbb{R}^n, \omega^{(t)} > 0 \} = \\ \text{Mínimo}_{\{U, \nu\}} \sum_{i=1}^T \omega^{(t)} \sum_{k=1}^K \sum_{j=1}^m (\mu_{ki})^\alpha \sum_{j=1}^n (X^{(t)}_{ij} - \nu^{(t)}_{kj})^2 | X_i, \nu_k \in \mathbb{R}^n, \omega^{(t)} > 0 \} = \\ \sum_{k=1}^K \sum_{i=1}^T \omega^{(t)} \{ \text{Mínimo}_{\{U, \nu\}} G(X_i, \nu_k) \}, \text{ donde } G(X_i, \nu_k) &\equiv \sum_i (\mu_{ki})^\alpha \sum_j (X^{(t)}_{ij} - \nu^{(t)}_{kj})^2 \end{aligned}$$

Los puntos estacionarios de $J(U, \nu)$ corresponden entonces a los de $G(X_i, \nu_k)$. De donde:

$$\partial G(X_i, \nu_k) / \partial \nu_k = 0 \Rightarrow \sum_i (\mu_{ki})^\alpha \sum_j (X^{(t)}_{ij} - \nu^{(t)}_{kj}) = 0 \Rightarrow \sum_j \nu^{(t)}_{kj} \sum_i (\mu_{ki})^\alpha = \sum_j \sum_i (\mu_{ki})^\alpha X^{(t)}_{ij}.$$

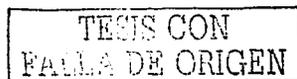
De donde se concluye la ecuación (2.22).

Para obtener (2.23), ahora suponemos que cada vector $\nu_k^{(t)}$ es fijo e independiente de i . Entonces, para $U=(\mu_{ki}) \in M_{bK}$:

$$\begin{aligned} \text{Mín}_{\{U\}} J(U, \nu) &= \text{Mín}_{\{U\}} \{ \sum_i \omega^{(t)} J^{(t)}(U, \nu_k^{(t)}) | \omega^{(t)} > 0, (\mu_{1i} + \mu_{2i} + \dots + \mu_{Ki}) = 1 \}, \\ &= \text{Mín}_{\{U\}} \{ \sum_i \omega^{(t)} \sum_k (\mu_{ki})^\alpha \sum_j (X^{(t)}_{ij} - \nu^{(t)}_{kj})^2 | \omega^{(t)} > 0, \sum_k (\mu_{ki}) = 1 \} \\ &= \sum_i \text{Mín} \{ H(\mu_{ki}) \}, \text{ donde } H(\mu_{ki}) \equiv \sum_i \omega^{(t)} \sum_k (\mu_{ki})^\alpha \sum_j (X^{(t)}_{ij} - \nu^{(t)}_{kj})^2 \end{aligned}$$

Por lo tanto, resolver el problema: $\text{Mín}_{\{U\}} \{ J(U, \nu) \}$ es equivalente a resolver el problema de optimización no lineal:

$$\text{Minimizar } H(\mu_{ki}) \equiv \sum_i \omega^{(t)} \sum_k (\mu_{ki})^\alpha \sum_j (X^{(t)}_{ij} - \nu^{(t)}_{kj})^2, \text{ bajo la restricción } \sum_k \mu_{ki} = 1.$$



Resolveremos este programa no lineal construyendo la función de Lagrange asociada:

$$L(\lambda, U) = H(\mu_{ki}) - \lambda [(\sum_k \mu_{ki}) - 1] = \sum_i \omega^{(i)} \sum_k (\mu_{ki})^\alpha \sum_j (x^{(i)}_{ij} - v^{(i)}_{kj})^2.$$

Se tiene:

$$(\partial L / \partial \lambda) = 0 \Rightarrow \sum_k \mu_{ki} = 1, \quad (2.24)$$

$$(\partial L / \partial U) = 0 \Rightarrow \sum_i \omega^{(i)} \sum_k (\mu_{ki})^{\alpha-1} \sum_j (x^{(i)}_{ij} - v^{(i)}_{kj})^2 = \sum_k (\lambda / \alpha) \Rightarrow (\lambda / \alpha) = (\mu_{ki})^{\alpha-1} \sum_i \omega^{(i)} \sum_j (x^{(i)}_{ij} - v^{(i)}_{kj})^2,$$

$$\text{de donde: } \mu_{ki} = (\lambda / \alpha)^{1/(\alpha-1)} \{ (\sum_i \omega^{(i)} \sum_j (x^{(i)}_{ij} - v^{(i)}_{kj})^2)^{1/(\alpha-1)} \}^{-1} \quad (2.25)$$

Aplicando la ecuación (2.24) a la ecuación anterior (2.25), ésta se convierte en:

$$(\lambda / \alpha)^{1/(\alpha-1)} \sum_q \{ \sum_i \omega^{(i)} \sum_j (x^{(i)}_{ij} - v^{(i)}_{kj})^2 \}^{1/(\alpha-1)} \}^{-1}; \quad q=1, 2, \dots, K;$$

$$\Rightarrow (\lambda / \alpha)^{1/(\alpha-1)} = (\sum_q \{ 1 / \sum_i \omega^{(i)} [d(x^{(i)}_{ij}, v^{(i)}_{kj})]^2 \}^{1/(\alpha-1)})^{-1}, \quad q=1, 2, \dots, K.$$

Sustituyendo este resultado en (2.25) obtenemos (2.23)■

Hemos demostrado entonces que la solución dada por las ecuaciones (2.22)-(2.23) es la que minimiza el valor de la función objetivo no lineal $J(U, v)$ en (2.21), es decir que (2.22)-(2.23) constituyen una SE del POEM (2.20).

Respecto al procedimiento algorítmico para la clasificación borrosa de D3F, que plantearemos en el siguiente capítulo de aplicaciones, es conveniente enfatizar que deberá estar basado en la aplicación inicial del algoritmo de clasificación de K-medias borrosas para D2F planteado anteriormente. Esto significa que, dado el conjunto de D3F con m objetos, n atributos y T situaciones; se parte del conjunto de datos correspondientes a $t=1$, a los cuales se les aplica dicho algoritmo; una vez alcanzado el criterio de convergencia para este conjunto inicial de datos, se repite el procedimiento para los datos correspondientes a $t=2$, partiendo precisamente de la clasificación de Pareto para la cual se alcanza el criterio de convergencia para $t=1$; y así sucesivamente, hasta que se haya cubierto todo el conjunto de D3F disponible. Una vez completado este proceso, el algoritmo de clasificación borrosa de D3F propuesto en el Teorema 3 nos permitirá determinar la clasificación de Pareto asociada al conjunto de D3F cuyas clases sean independientes de las situaciones y que, por lo tanto, constituyan una representación global válida de los D3F a lo largo de las diversas situaciones incluidas en el estudio.

Capítulo 3. Aplicaciones del modelo de clasificación: El patrón de evolución económica de México en el contexto de la OCDE.

TESIS CON
FALLA DE ORIGEN

3.1 Presentación

En este capítulo se pretende establecer el patrón de evolución de la economía mexicana en el contexto del desarrollo capitalista de los países que integran la Organización para la Cooperación y el Desarrollo Económicos (OCDE). El período de estudio comprende, prácticamente, la totalidad de la segunda mitad del siglo XX, específicamente desde 1950 hasta 1992. Dada la naturaleza del presente trabajo, descrita previamente en la sección 2.1, el principal objetivo es revelar (o, al menos, contribuir a dilucidar) la estructura económica latente de los datos de 3 formas (D3F) reportados en las fuentes documentales de información económica conocidas y, sobre todo, reconocidas respecto a la situación de la economía mexicana. Desde una perspectiva estrictamente económica, es posible formular tres objetivos específicos del presente estudio:

- Lograr una descripción global de la economía mundial válida para el mayor período histórico posible.
- Proponer un modelo matemático que contribuya a identificar, de manera sistemática, las tendencias de convergencia y/o divergencia de los niveles de vida en las economías de los países incluidos en el estudio.
- Aportar elementos de juicio fundamentados para dilucidar las características distintivas de las fases de desarrollo del crecimiento económico de los países en general y de nuestro país en particular.

En la medida que este trabajo se acerque a lograr estos objetivos, podremos afirmar la posible pertinencia del mismo.

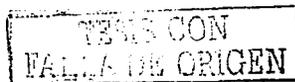
El primer problema al que nos enfrentamos es localizar fuentes de información económica de reconocida calidad y objetividad. Para solventar este problema se tomó la decisión de emplear los datos económicos publicados en la propia OCDE, Maddison [1997]. Aunque en esta referencia se afirma que constituye la más completa base de datos disponible para el análisis cuantitativo y comparativo del desempeño económico de los países (el autor asevera que se incluyen estadísticas de la economía mundial en el período 1820-1992), en realidad se reportan muchos vacíos de información, indicados con la notación "n.d.". Por lo tanto, para poder emplear información económica realmente documentada, y que incluyera a México, se redujo el conjunto de atributos documentados a las siguientes cuatro variables económicas:

- X_1 : Años de educación formal en la población de 15 a 64 años
- X_2 : Nivel del PIB per cápita
- X_3 : Productividad del trabajo (PIB por hora trabajada)
- X_4 : PIB por persona empleada

Considerando este universo de 4 atributos, el número de países de estudio con información documental completa al respecto, es de sólo 20; y únicamente para tres años: 1950, 1973 y 1992. Considerando estas restricciones documentales, el conjunto de D3F considerado en el presente estudio consiste en la matriz de datos:

$$X^{(i)}_{ij}, \text{ donde } 1 \leq i \leq 20, 1 \leq j \leq 4, 1 \leq t \leq 3 \quad (3.1)$$

En la siguiente sección establecemos el algoritmo de clasificación borrosa para los D3F dados en (3.1), el cual aplicaremos luego al caso reducido de sólo dos atributos y, posteriormente, al universo completo de los cuatro atributos mencionados.



3.2 El algoritmo de clasificación borrosa para D3F

Paso 1. Establecer: el número K apropiado de clases, el criterio de convergencia ϵ_L , $\epsilon_L > 0$, y el valor del nivel de borrosidad α , para el conjunto de D3F dado en (3.1). Aplicar el algoritmo de clasificación para K -medias borrosas, descrito en la sección 2.4, al conjunto de datos (3.1) correspondientes a $t=1$ hasta que se alcance el criterio de convergencia ϵ_L previamente establecido.

Paso 2. Emplear como solución inicial los valores de pertenencia μ_{ki} , obtenidos en el paso 1, al conjunto de datos (3.1) correspondientes a $t=2$ hasta que se alcance, nuevamente, el criterio de convergencia ϵ_L .

Paso 3. Repetir el paso 2 hasta que éste se haya completado para todas las T situaciones.

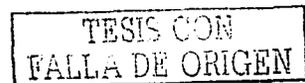
Paso 4. Determinar el conjunto de parámetros de ponderación $\{\omega_i\}$ que se utilizará para aplicar el Teorema 3. Emplear los valores de las distancias euclidianas $d_{ki} = (d_{i,v}^{(t)}, v^{(t)_k})$, asociadas a las T soluciones de convergencia ya obtenidas, con los valores de los parámetros (ϵ, α) previamente utilizados, para determinar los valores de pertenencia (μ_{ki}) del objeto i a la clase borrosa k , de acuerdo al Teorema 3.

Paso 5. Dados los valores (μ_{ki}) obtenidos, el objeto o_i se asigna a la clase C_k cuando $\mu_{ki} > \mu_{pi}$, $\forall k, p \in \{1, 2, \dots, K\}$ con $k \neq p$. Se logra así una representación en R^n de las K clases al calcular los respectivos centroides de las clases obtenidas empleando el Teorema 3.

En el paso 1 se recomienda que el número apropiado de clases surja como resultado de una investigación conducida por (o como una propuesta de) expertos en el campo específico de conocimiento de los D3F. De otra manera es posible que se llegue a situaciones de plantear escenarios del tipo "y qué pasa si", o bien, del tipo ensayo y error, para proponer el valor de K . Sin embargo, la mayoría de los estudios económicos serios emplean actualmente una clasificación (borrosa) de orden $K=3$: *países altamente industrializados, semi-industrializados y pobremente industrializados*. En consecuencia, utilizaremos en este trabajo el valor $K=3$ y mantendremos este valor en la medida en que nos conduzca a resultados lógicamente aceptables.

Respecto al valor $\alpha=2$ que emplearemos, consideramos que nos debe permitir verificar la convergencia hacia un valor mínimo óptimo, o ideal, de acuerdo con (2.7); de no ser así, se recomienda realizar un estudio heurístico (básicamente tomando valores alrededor de 2) para determinar cuál valor α permite una mejor clasificación en el número K de clases previamente propuesto. Es posible reducir la escala del estudio de los cuatro atributos definidos $\{X_1, X_2, X_3, X_4\}$, a sólo dos, pues podemos postular al atributo X_2 como representativo de $\{X_2, X_3, X_4\}$, ya que se trata de la misma variable sólo que mide sobre rangos distintos.

Finalmente, en el paso 4, se aplica el criterio de Laplace empleado en la toma de decisiones en escenarios probabilísticos, para determinar equitativamente los factores de ponderación $\{\omega_i\}$, pues desconocemos cuál, o cuáles, de las $T=3$ situaciones influye más decisivamente en los valores observados de los $N=4$ atributos considerados. Si se obtuviera semejante información podríamos utilizar un criterio de ponderación discriminatorio o preferencial



para enfatizar aquellas situaciones más influyentes y, recíprocamente, para disminuir las situaciones menos influyentes. Incluso podría entonces realizarse un análisis visual a escala reducida, de hasta tres atributos o variables de estudio, (usando en tal caso una gráfica tridimensional) para determinar cuál valor de ω_i , con $1 \leq i \leq 3$ y $\sum_i \omega_i = 1$, es el que permite distinguir más claramente las K clases formadas.

3.3 Estudio de clasificación a escala reducida

En esta sección realizaremos el estudio a escala reducida para aplicar el algoritmo de clasificación borrosa, introducido en la sección 3.2, a los D3F definidos en (3.1) considerando sólo dos de las cuatro variables económicas definidas.

Considerando que X_2 , X_3 y X_4 definen a la misma variable PIB, aunque sobre rangos distintos de las poblaciones de cada país (ver sección 3.1), entonces basta con estudiar el comportamiento concomitante del subconjunto dado por el par de atributos $\{X_1, X_2\}$; la matriz de D3F para este subconjunto completo de atributos se muestra en la Tabla 6.

PAIS	Objeto	Datos para 1950		Datos para 1973		Datos para 1992	
		Xi1	Xi2	Xi1	Xi2	Xi1	Xi2
Bélgica	1	9.83	5346	11.99	11905	15.24	17165
Francia	2	9.58	5221	11.69	12940	15.96	17959
Alemania	3	10.4	4281	11.55	13152	12.17	19351
Italia	4	5.49	3425	7.62	10409	11.2	16229
P. Bajos	5	8.12	5850	10.27	12763	13.34	16898
Suecia	6	9.5	6738	10.44	13494	14.24	16927
Reino U.	7	10.84	6847	11.66	11992	14.09	15738
Portugal	8	2.53	2132	4.62	7568	9.11	11130
España	9	5.13	2397	6.29	8739	11.51	12498
E.U.	10	11.27	9573	14.58	16607	18.04	21558
Argent.	11	4.8	4987	7.04	7970	10.7	7616
Brasil	12	2.05	1673	3.77	3913	6.41	4637
Chile	13	5.47	3827	7.98	5028	10.93	7238
Colomb.	14	2.66	2089	4.91	3539	9.14	5025
México	15	2.6	2085	5.22	4189	8.22	5112
Venez.	16	2.21	7424	4.41	10717	10.18	9163
India	17	1.35	597	2.6	853	5.55	1348
Japón	18	9.11	1873	12.09	11017	14.86	19425
Corea	19	3.36	876	6.82	2840	13.55	10010
Taiwán	20	3.62	922	7.35	3669	13.83	11590
Promedio		5.996	3908.15	8.145	8665.2	11.9135	12330.9
Desv. Est.		3.36287	2454.44	3.32184	4357.05	3.14921	5742.37

Tabla 6. Subconjunto de D3F reportados de $X_{ij}^{(t)}$ para $j=1$ (Años de educación formal en la población de 15 a 64 años) y $j=2$ (PIB per cápita), correspondientes a los tiempos $t=1$ (1950), $t=2$ (1973) y $t=3$ (1992) para 20 países miembros de la OCDE [Maddison, 1997].

En este caso la clasificación se considera como un problema de optimización de tres criterios, es decir, $T=3$ en (2.20). Puesto que los valores de los tres criterios deben ser comparables, trabajaremos con los datos estandarizados: media 0 y varianza 1, con respecto a cada variable para cada tiempo. En la parte inicial consideremos los respectivos diagramas de dispersión para los D3F estandarizados de la Tabla 5: *PIB per cápita* versus *Años de educación formal*, para cada uno de los tres años mencionados que se muestran, por separado para propósitos estrictamente comparativos, en las figuras (que, por razones de espacio de presentación hemos llamado gráficas) 3, 4 y 5.

Gráfica 3: PIB per cápita versus Años de educación formal para D3F estandarizados (1950)

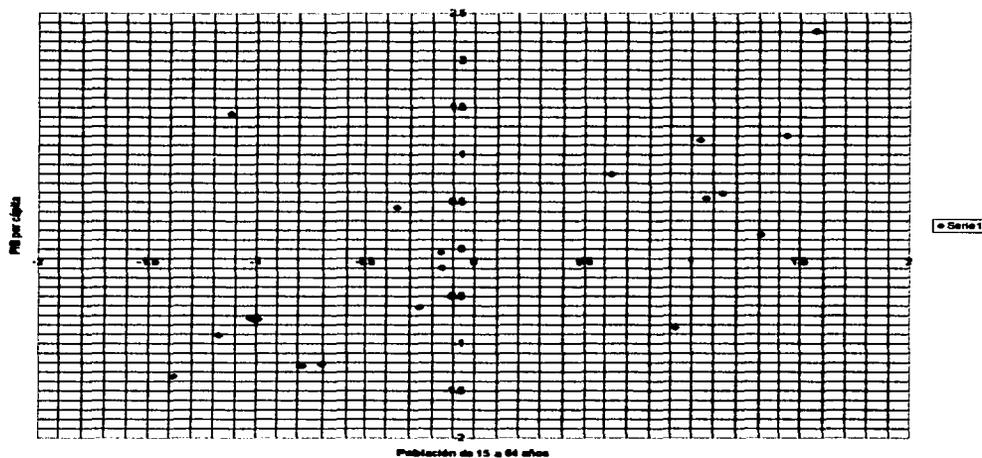
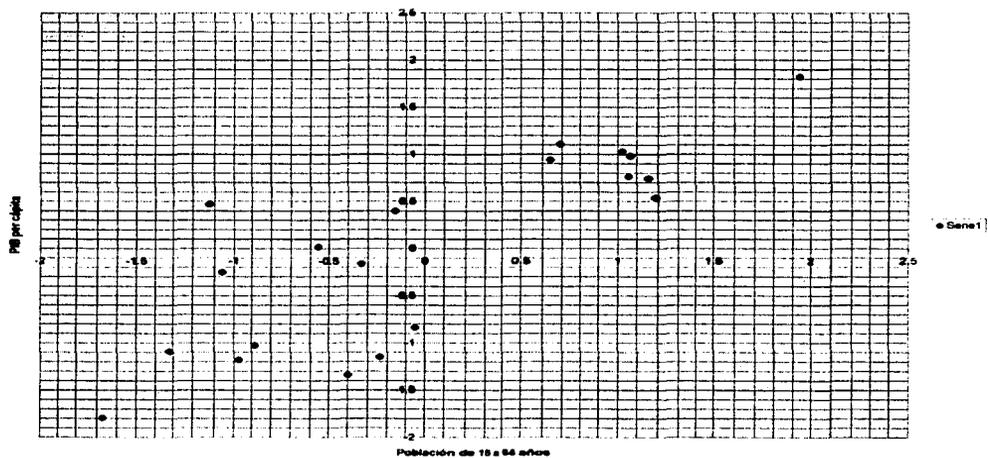
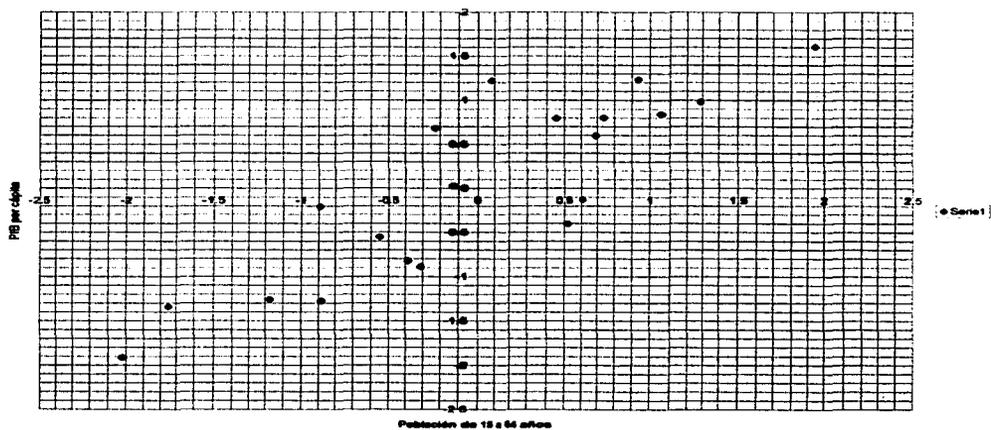


Figura 3. *PIB per cápita* versus *Años de educación formal* para los D3F estandarizados dados en la Tabla 6 correspondientes a 1950.

Gráfica 4: PIB per cápita vs. Años de educación formal para D3F estandarizados (1973)



Gráfica 5: PIB per cápita vs. Años de educación formal para D3F estandarizados (1992)



TESIS CON
FALLA DE ORIGEN

Los diagramas de dispersión muestran claramente la tendencia de los objetos de estudio a agruparse en tres clases, lo cual es excepcionalmente claro en el caso correspondiente a $t=2$ (1973). Esto nos alienta a mantener la propuesta inicial de $K=3$ para el modelo de clasificación borrosa global.

Por otra parte, como no se tiene información acerca de cuál situación temporal es más importante desde el punto de vista económico, asignaremos ponderaciones $\omega^{(t)}$ iguales para todas las funciones objetivo $J^{(t)}$ en (2.20).

Los valores iniciales de pertenencia $U=(\mu_{ki})$ se asignan aleatoriamente, y se determina un criterio de convergencia $\epsilon_L=0.01$ para detener el procedimiento algorítmico. Los resultados se obtienen en 8, 6 y 7 iteraciones, para los datos de 1950, 1973 y 1992, respectivamente. Las Tablas 7, 8, 9 y 10 muestran los valores de las matrices de pertenencia borrosa $U=(\mu_{ki})$ obtenidos para cada uno de los veinte países incluidos en el estudio, tanto para los tres años indicados como para el patrón global de evolución económica. En la Gráfica 6 se muestra el patrón global de evolución económica, los datos estandarizados para 1992, así como los correspondientes centroides de las tres clases construidas. Similarmente, en la Tabla 11 se muestran los valores de los centroides característicos de cada clase para cada uno de los tres años, así como los centroides asociados al patrón global de evolución económica, mostrados en la Gráfica 6.

Podemos observar en las Gráficas 3-5 cómo se distinguen claramente las tres clases en este estudio restringido a sólo dos de las cuatro variables económicas.

Este hecho es confirmado al examinar las Tablas 7-9 donde se muestran para cada país su pertenencia borrosa a cada clase a la cual sospechamos que pertenece. En el caso específico de México, se observa que prácticamente coincide con Colombia y Portugal en 1950; en 1973 Portugal se ha separado de México y Colombia (al pasar a la clase 2), quienes se mantienen con valores cercanos entre sí; para 1992 Portugal se ha despegado notablemente de México, cuya situación económica ha empeorado visiblemente en ambas variables económicas durante este lapso.

En particular, para los resultados en la Tabla 7, se pueden identificar los países que integran las tres clases borrosas donde, obviamente, los países con mejores niveles de los indicadores económicos pertenecen a la clase 1 y aquellos con los peores niveles de vida que conforman la clase 3. Se muestran en negritas a los líderes de cada clase, es decir, aquellos países que poseen la más alta pertenencia a su respectiva clase.

Las correspondientes tablas análogas para los años 1973 y 1992 nos permitirán estudiar la evolución temporal de las clases formadas. Esto, a su vez, nos ayudará a comprender mejor el modelo global que describe la evolución económica de los veinte países. Todo con el fin de lograr una explicación objetiva de la evolución económica específica de México en el contexto global durante el lapso 1950-1992.

TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN

PAIS	Objeto	Datos (1950)		Matriz U borrosa		
		Xi1	Xi2	U1i	U2i	U3i
Bélgica	1	1.1401	0.58582	0.99667	0.00303	0.0003
Francia	2	1.06576	0.53489	0.99199	0.00736	0.00066
Alemania	3	1.3096	0.15191	0.93143	0.0598	0.00877
Italia	4	-0.1505	-0.1968	0.0019	0.98782	0.01028
Países B.	5	0.6316	0.79116	0.94165	0.05577	0.00258
Suecia	6	1.04197	1.15295	0.99942	0.00053	5.3E-05
Reino U.	7	1.44044	1.19736	0.99842	0.00138	0.0002
Portugal	8	-1.0307	-0.7236	7.2E-05	0.00229	0.99764
España	9	-0.2575	-0.6157	0.00953	0.54908	0.4414
E.U.	10	1.5683	2.308	0.926	0.06076	0.01324
Argentina	11	-0.3556	0.43955	0.00137	0.99708	0.00156
Brasil	12	-1.1734	-0.9107	2.2E-05	0.00046	0.99952
Chile	13	-0.1564	-0.0331	0.00028	0.99891	0.00081
Colombia	14	-0.992	-0.7412	5.1E-05	0.00164	0.99831
México	15	-1.01	-0.743	0.00005	0.00157	0.99838
Venezuela	16	-1.1258	1.43245	0.14029	0.73478	0.12494
India	17	-1.3816	-1.349	0.00068	0.00764	0.99169
Japón	18	0.926	-0.8292	0.28482	0.50219	0.21299
Corea	19	-0.7839	-1.2354	0.00016	0.00245	0.99739
Taiwán	20	-0.7065	-1.2166	0.00027	0.00428	0.99544

Tabla 7. Datos estandarizados para las variables $X^{(1950)}_{i1}$: Años de educación formal en la población de 15 a 64 años, y $X^{(1950)}_{i2}$: PIB per cápita; y los respectivos niveles de pertenencia (μ_{ki}) al aplicar el paso 1 del Algoritmo de Clasificación dado en 3.2, con un criterio de convergencia $\epsilon_L \leq 0.1$, para 20 países miembros de la OCDE.

De acuerdo a la Tabla 7, los países que integran a la clase borrosa C_k , $k=1,2,3$; pues su nivel respectivo de pertenencia borrosa máxima ocurre precisamente respecto a C_k , son:

Clase 1	Clase 2	Clase 3
Bélgica	Italia	Portugal
Francia	España	Brasil
Alemania	Argentina	Colombia
Países Bajos	Chile	México
Suecia	Venezuela	India
Reino Unido	Japón	Corea
Estados Unidos		Taiwán

Donde, como se mencionó previamente, los líderes de la clase k , es decir aquellos países cuyo nivel de pertenencia en la clase k es el máximo, se muestran en negritas.

PAIS	Objeto	Datos (1973)		Matriz U borrosa		
		Xi1	Xi2	U1i	U2i	U3i
Bélgica	1	1.15749	0.74358	0.99963	0.00031	6.3E-05
Francia	2	1.06718	0.98112	1	2.4E-08	0
Alemania	3	1.02504	1.02978	1	3E-06	1E-06
Italia	4	-0.158	0.40023	0.0273	0.96259	0.01012
Países B.	5	0.63971	0.9405	0.99148	0.00765	0.00087
Suecia	6	0.69088	1.10827	0.99559	0.00392	0.0005
Reino U.	7	1.05815	0.76354	0.99972	0.00023	4.4E-05
Portugal	8	-1.0612	-0.2518	0.00217	0.91376	0.08407
España	9	-0.5584	0.01694	4.9E-06	0.99997	2.5E-05
E.U.	10	1.93718	1.82275	0.9721	0.0212	0.0067
Argentina	11	-0.3326	-0.1596	0.00178	0.98819	0.01003
Brasil	12	-1.317	-1.0907	0.00052	0.01568	0.98379
Chile	13	-0.0497	-0.8348	0.01994	0.31306	0.667
Colombia	14	-0.9739	-1.1765	3.6E-06	1E-04	0.9999
México	15	-0.8805	-1.0273	3E-05	0.001	0.999
Venezuela	16	-1.1244	0.47092	0.00596	0.97574	0.01831
India	17	-1.6693	-1.793	0.00396	0.04404	0.95201
Japón	18	1.1876	0.53977	0.99565	0.00354	0.00082
Corea	19	-0.3989	-1.337	0.00082	0.0112	0.98799
Taiwán	20	-0.2393	-1.1467	0.00347	0.05091	0.94562

Tabla 8. Datos estandarizados para las variables $X^{(1973)}_{i1}$: Años de educación formal en la población de 15 a 64 años, y $X^{(1973)}_{i2}$: PIB per cápita; y los respectivos niveles de pertenencia (μ_{ki}) al aplicar el paso 2 del Algoritmo de Clasificación dado en 3.2, con un criterio de convergencia $\epsilon_L \leq 0.1$, para 20 países miembros de la OCDE.

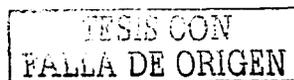
Para los datos de 1973, los países que integran cada una de las tres clases borrosas son:

Clase 1	Clase 2	Clase 3
Bélgica	Italia	Brasil
Francia	Portugal	Chile
Alemania	España	Colombia
Países Bajos	Argentina	México
Suecia	Venezuela	India
Reino Unido		Corea
Estados Unidos		Taiwán
Japón		

Como puede observarse, ocurren tres transiciones en el status 1973 respecto al de 1950:

1. Japón pasa de la clase 2 a la clase 1, mejorando sustancialmente ambas variables.
2. Portugal pasa de la clase 3 a la clase 2, al mejorar a más del doble ambas variables.
3. Chile pasa de la clase 2 a la clase 3, pues no aumenta suficientemente sus variables.

Es notable que los países de la clase 1 mantienen su status sin modificaciones en este lapso.



PAIS	Objeto	Datos para 1992		Matriz U borrosa		
		Xi1	Xi2			
Bélgica	1	1.0563	0.84184	0.9997	0.00028	2E-05
Francia	2	1.28493	0.98011	0.99825	0.0016	0.00015
Alemania	3	0.08145	1.22252	0.93284	0.06228	0.00488
Italia	4	-0.2266	0.67884	0.41835	0.56207	0.01958
Países B.	5	0.45297	0.79534	0.98542	0.01397	0.00061
Suecia	6	0.73876	0.80039	0.99941	0.00056	3E-05
Reino U.	7	0.69113	0.59334	0.98809	0.01146	0.00046
Portugal	8	-0.8902	-0.2091	0.01085	0.92547	0.06368
España	9	-0.1281	0.02911	0.00861	0.98938	0.00201
E.U.	10	1.9454	1.6069	0.96504	0.02994	0.00502
Argentina	11	-0.3853	-0.8211	0.00197	0.97951	0.01852
Brasil	12	-1.7476	-1.3398	4.7E-05	0.00067	0.99928
Chile	13	-0.3123	-0.8869	0.00269	0.97392	0.02339
Colombia	14	-0.8807	-1.2723	0.00203	0.09005	0.90792
México	15	-1.173	-1.257	0.0002	0.00594	0.99386
Venezuela	16	-0.5505	-0.5517	0.00103	0.9917	0.00726
India	17	-2.0207	-1.9126	0.00091	0.00818	0.99091
Japón	18	0.93563	1.23541	0.99971	0.00027	2.4E-05
Corea	19	0.51965	-0.4042	0.06171	0.92808	0.01021
Taiwán	20	0.60857	-0.129	0.23502	0.75395	0.01103

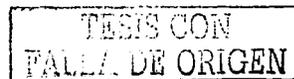
Tabla 9. Datos estandarizados para las variables $X^{(1992)}_{11}$: Años de educación formal en la población de 15 a 64 años, y $X^{(1992)}_{12}$: PIB per cápita; y los respectivos niveles de pertenencia (μ_{ki}) al aplicar el paso 3 del Algoritmo de Clasificación dado en 3.2, con un criterio de convergencia $\epsilon_i \leq 0.1$, para 20 países miembros de la OCDE.

En este caso, las clases identificadas y los respectivos países que las forman son:

Clase 1	Clase 2	Clase 3
Bélgica	Italia	Brasil
Francia	Portugal	Colombia
Alemania	España	México
Países Bajos	Argentina	India
Suecia	Chile	
Reino Unido	Venezuela	
Estados Unidos	Corea	
Japón	Taiwán	

Como puede observarse, esta clasificación (1992) difiere de la obtenida en el caso anterior (1973) en cinco hechos fundamentales:

1. Chile regresa de la clase 3 a la clase 2 (como en 1950), al mejorar ambas variables.
2. Corea pasa de la clase 3 a la clase 2: $X^{(1992)}_{19,2}$ es 2.5 veces más que $X^{(1973)}_{19,2}$.
3. Similarmente, Taiwán pasa de la clase 3 a la clase 2: $X^{(1992)}_{20,2} = 2.16 (X^{(1973)}_{20,2})$
4. Finalmente, Italia está cerca de pasar de la clase 2 a la clase 1.
5. Japón se convierte en el nuevo líder de la clase 1.



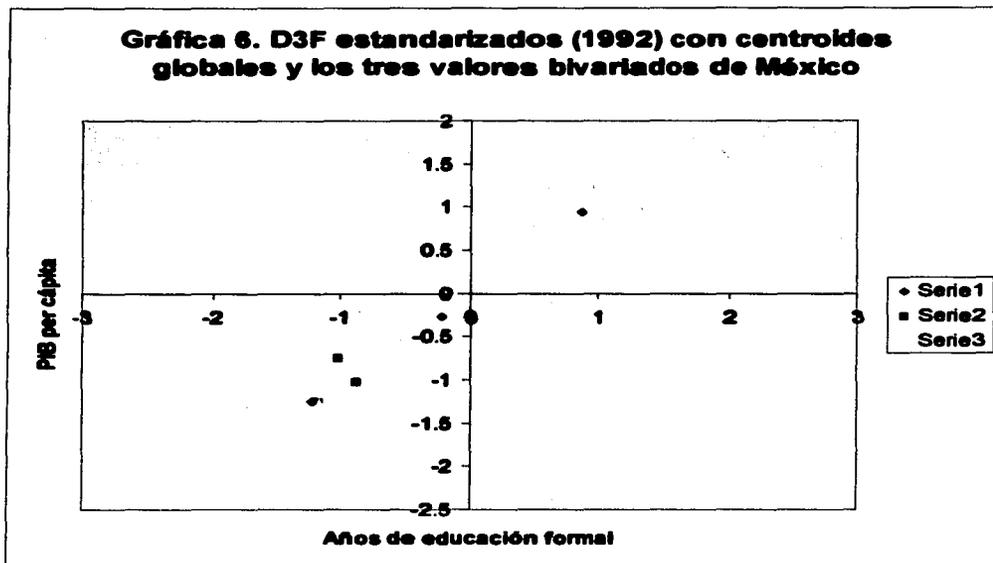
PAIS	Objeto	Matriz U borrosa		
		U1i	U2i	U3i
Bélgica	1	0.96586	0.02556	0.00859
Francia	2	0.95628	0.03262	0.0111
Alemania	3	0.82987	0.12545	0.04468
Italia	4	0.18186	0.71064	0.1075
Países B.	5	0.85652	0.11454	0.02895
Suecia	6	0.9538	0.03563	0.01057
Reino U.	7	0.94567	0.04065	0.01368
Portugal	8	0.06337	0.54281	0.39382
España	9	0.06412	0.83391	0.10197
E.U.	10	0.78096	0.14856	0.07048
Argentina	11	0.0375	0.88749	0.075
Brasil	12	0.01083	0.048	0.94117
Chile	13	0.081	0.67353	0.24547
Colombia	14	0.01655	0.09586	0.88759
México	15	0.0085	0.0488	0.9427
Venezuela	16	0.13208	0.67697	0.19096
India	17	0.03507	0.11239	0.85255
Japón	18	0.67015	0.23398	0.09587
Corea	19	0.12243	0.46672	0.41086
Taiwán	20	0.14438	0.51214	0.34347

Tabla 10. Niveles de pertenencia (μ_{ij}) para las variables $X^{(1)}_{11}$: Años de educación formal en la población de 15 a 64 años, y $X^{(1)}_{12}$: PIB per cápita; al aplicar el paso 4 del Algoritmo de Clasificación dado en 3.2, con un criterio de convergencia $\epsilon_L \leq 0.1$, para 20 países miembros de la OCDE. En negritas se muestran los países líderes de cada clase.

Para este caso global, obtenido al aplicar el paso 4 del algoritmo de clasificación borrosa dado en 3.2 a los datos correspondientes a 1992, las clases identificadas y los respectivos países que las integran son:

Clase 1	Clase 2	Clase 3
Bélgica	Italia	Brasil
Francia	Portugal	Colombia
Alemania	España	México
Países Bajos	Argentina	India
Suecia	Chile	
Reino Unido	Venezuela	
Estados Unidos	Corea	
Japón	Taiwán	

Esta es la representación global de la evolución económica de los 20 países incluidos en el presente estudio, la cual no necesariamente permite observar las características particulares detectadas en cada una de las situaciones temporales previamente descritas; por ejemplo, que en 1992 Italia se encuentra a punto de hacer la transición de la clase 2 a la clase 1.

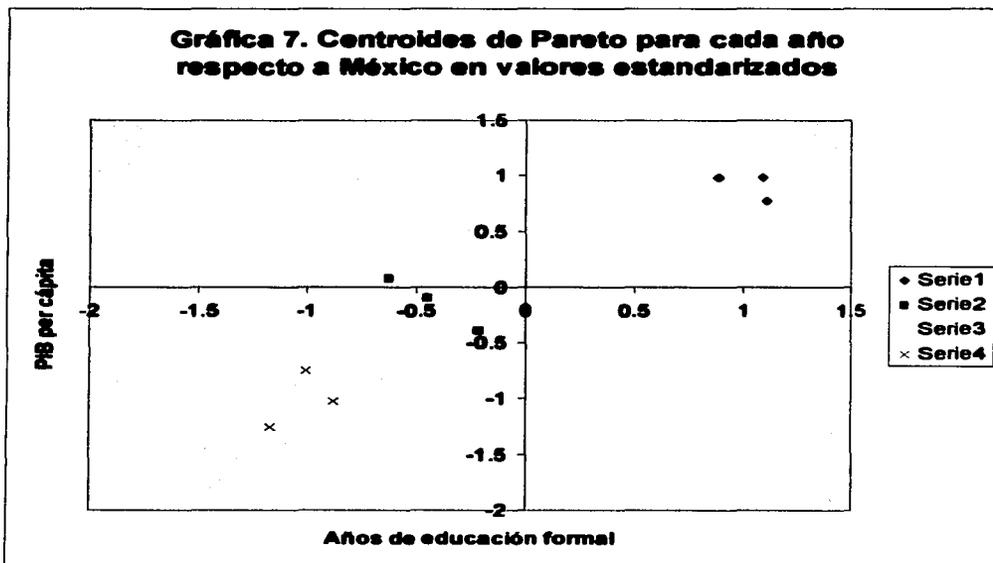


Gráfica 6. Datos estandarizados para el año 1992(), centroides globales (●) y los valores estandarizados correspondientes a México (□).

Considerando la Tabla 10 y la Gráfica 6 podemos destacar cuatro hechos importantes:

1. Los tres líderes han pertenecido a la misma clase durante todo el lapso 1950-1992.
2. Portugal, Corea y Taiwán de la clase 2 tienen pertenencia significativa en la clase 3.
3. Estados Unidos tiene la menor pertenencia (0.78096) en la clase 1. Esto pudiera deberse a que, en este primer estudio, sólo se incluyen dos variables: *Años de educación formal en la población de 15 a 64 años y PIB per cápita*.
4. En el caso específico de México observamos cómo se reducen sus valores asociados en ambas variables, al grado de que es el objeto más cercano al centroide global de la clase 3: (-1.2283, -1.2548), para el año 1992.
5. Es importante recalcar el punto 4 de la siguiente manera: Pese a que el patrón global de evolución es hacia el primer cuadrante, México presenta una tendencia en el sentido opuesto, es decir, hacia el tercer cuadrante.

Para verificar la consistencia de este análisis, en la Gráfica 7 se exhiben los centroides (dados en la Tabla 11) de los tres años junto con los valores estandarizados de México.



Gráfica 7. Centroides de las tres clases para cada año de los datos de la Tabla 6, junto con los valores estandarizados (x) correspondientes a México.

Nuevamente observamos que México muestra en este lapso una muy alta pertenencia a la clase 3 y, de hecho, globalmente se convierte en líder de esta clase, como se comprueba en la Tabla 10 y en la Gráfica 6.

Desde la perspectiva del universo reducido de las dos variables incluidas podemos concluir este estudio a escala reducida comentando respecto al logro de los objetivos específicos planteados inicialmente en la sección 3.1:

1. Respecto al objetivo de lograr una descripción global válida de la economía mundial.
 - Consideramos que se ha logrado una descripción global adecuada de la economía mundial durante el período de estudio. De acuerdo a la Tabla 11: el PIB per cápita en las clases 1 y 2 aumentó primero rápido (de 1950 a 1973) y luego lentamente (de 1973 a 1992); pero en la clase 3 ha ido reduciéndose drásticamente de 1950 a 1992. En cambio, la variable Años de educación formal se ha reducido, paulatinamente para todas las clases (con una leve “mejoría” para las clases 2 y 3 durante el período 1973-1992)
 - Los líderes de cada clase han pertenecido a su respectiva clase a lo largo de todo el período de estudio.

Clase	Año	X ₁	X ₂
1	1950	1.11465	0.77654
	1973	1.09068	0.98574
	1992	0.8795	0.98785
	GLOBAL	0.87213	0.94504
2	1950	-0.4444	-0.0909
	1973	-0.624	0.0788
	1992	-0.2103	-0.3877
	GLOBAL	-0.2165	-0.2697
3	1950	-1.0677	-0.8898
	1973	-0.8531	-1.2252
	1992	-1.4778	-1.4497
	GLOBAL	-1.2283	-1.2548

Tabla 11. Centroides temporales y globales para los datos de la Tabla 6.

- Las transiciones interclase observadas han ocurrido en aquellos países donde se tienen niveles comparativamente altos o bajos en, al menos, una de las variables. Quizá el ejemplo más sorprendente lo constituye Japón, quién de un nivel de pertenencia a la clase 1 de .28482 en 1950, pasó a .99565 en 1973, y a .99971 en 1992.
 - La representación global permite identificar las clases globales así como los líderes correspondientes en cada clase, y, lamentablemente, permite clasificar a México como líder global de la clase 3.
2. Se ha propuesto un modelo matemático que permite identificar sistemáticamente las posibles tendencias de convergencia y/o divergencia de los niveles observados en los países incluidos en el estudio, específicamente se ha comentado el caso de Japón y el de México. A este respecto, México muestra una tendencia divergente respecto a los niveles observados en los países pertenecientes a la clase 1, como se observa en la Gráfica 7.
 3. Consideramos que este estudio inicial a escala reducida contiene y, por lo tanto, puede aportar elementos de juicio económico y social tendientes a establecer las características que distinguen las diversas fases de desarrollo del crecimiento económico de los países.
 4. Para fines comparativos y/o de validación de los resultados obtenidos en el presente estudio, puede calcularse la medida de bondad de clasificación (2.14) empleando los valores reportados en las tablas 10 y 11.

En la siguiente sección aplicaremos este análisis restringido al conjunto completo de D3F.

TESIS CON
FALLA DE ORIGEN

3.4 Clasificación económica de México en el contexto de la OCDE

En esta parte realizaremos el estudio completo sobre la clasificación económica de México en el contexto de la OCDE considerando las cuatro variables económicas definidas previamente en la sección 3.1:

- X₁: Años de educación formal en la población de 15 a 64 años
- X₂: Nivel del PIB per cápita
- X₃: Productividad del trabajo (PIB por hora trabajada)
- X₄: PIB por persona empleada.

De modo que el universo de datos de estudio es el conjunto de D3F dado en (3.1), cuyas $20 \times 4 \times 3 = 240$ observaciones se muestran en la Tabla 12. Como se mencionó en el estudio a escala reducida realizado en la sección anterior, los datos son estandarizados para propósitos comparativos empleando los promedios y desviaciones estándar reportados al pie de la Tabla 12.

Resultados obtenidos

El procedimiento seguido en este caso es el mismo ya descrito en el estudio a escala reducida realizado en la sección 3.3, con la excepción de que no es posible ahora explicar los resultados obtenidos en este estudio completo por medio de un análisis gráfico de los D3F estandarizados. Al aplicar el algoritmo de clasificación de K-medias borrosas a cada una de las tres situaciones temporales, se necesitaron de 13, 11 y 8 iteraciones para los años 1950, 1973 y 1992, respectivamente. Esto es consistente con el hecho de que, al aplicar el algoritmo en la situación temporal t , se emplea la solución convergente alcanzada en la situación temporal $(t-1)$, por lo que, en general, se requieren menos iteraciones al considerar los sucesivos conjuntos de datos temporales. Prescindiendo del apoyo gráfico, se muestran en la Tabla 13 los niveles de pertenencia borrosa obtenidos para cada uno de las tres situaciones temporales consideradas en el presente trabajo. En la Tabla 14 se muestran las distancias a los centroides de las tres clases identificadas, siguiendo el procedimiento descrito en el estudio a escala reducida de la sección anterior, para cada uno de los tres años considerados. En particular podemos observar en la Tabla 13 la continua pertenencia de México a la clase 3 durante todo el lapso histórico con un nivel notablemente alto, ≥ 0.99845 . Un nivel de pertenencia también muy alto se observa para México en la Tabla 19, donde aparecen los elementos de la matriz borrosa U para el modelo global, al considerar a los datos como D3F para los cuales se alcanza el criterio de convergencia $\epsilon_{i,} \leq 0.01$; en esta representación global México es prácticamente el líder de la clase 3, pues dista sólo 0.00798 del nivel máximo que ostenta Colombia.

Es importante analizar el hecho observado en la Tabla 19, donde se muestran los niveles de pertenencia obtenidos para los D3F globales, en esta tabla Estados Unidos aparece como el único integrante de la clase 2, lo cual no fue observado durante la realización del estudio a escala reducida de la sección anterior. Sin embargo sí ocurre al incluir al conjunto total de cuatro variables en el procedimiento algorítmico descrito.

VARIABLES DE ESTUDIO: DATOS DE 3 FORMAS

País	Variables de estudio 1950, (t=1)				Variables de estudio 1973, (t=2)				Variables de estudio 1992, (t=3)			
	Xi1	Xi2	Xi3	Xi4	Xi1	Xi2	Xi3	Xi4	Xi1	Xi2	Xi3	Xi4
Objeto												
Bélgica	9.83	5346	6.06	13826	11.99	11905	16.53	30943	15.24	17165	28.55	45260
Francia	9.58	5221	5.65	11108	11.69	12940	17.77	31464	15.96	17959	29.62	45678
Alemania	10.4	4281	4.37	10110	11.55	13152	16.64	30012	12.17	19351	27.55	43061
Italia	5.49	3425	4.28	8548	7.62	10409	15.58	25110	11.2	16229	24.59	36632
Países Bajos	8.12	5850	6.5	14361	10.27	12763	19.02	33304	13.34	16898	28.8	38538
Suecia	9.5	6738	7.08	13814	10.44	13494	18.02	28307	14.24	16927	23.11	35016
Reino Unido	10.84	6847	7.86	15395	11.66	11992	15.92	26882	14.09	15738	23.98	35751
Portugal	2.53	2132	2.58	5678	4.62	7568	9.86	18736	9.11	11130	14.06	23900
España	5.13	2397	2.6	5727	6.29	8739	10.86	23346	11.51	12498	20.22	38639
Estados Unidos	11.27	9573	12.66	23643	14.58	16607	23.45	40526	18.04	21558	29.1	46242
Argentina	4.8	4987	6.16	12538	7.04	7970	10.7	21349	10.7	7616	11.86	21666
Brasil	2.05	1673	2.41	4922	3.77	3913	5.62	11781	6.41	4637	6.66	12366
Chile	5.47	3827	4.66	10316	7.98	5028	8.9	17404	10.93	7238	10.66	21375
Colombia	2.66	2089	2.79	6492	4.91	3539	5.87	12578	9.14	5025	7.76	15124
México	2.6	2085	3.09	6665	5.22	4189	7.63	15728	8.22	5112	8.4	17327
Venezuela	2.21	7424	9.01	23791	4.41	10717	19.31	37936	10.18	9163	16.73	31255
India	1.35	597	0.6	1328	2.6	853	0.94	2064	5.55	1348	1.58	3482
Japón	9.11	1873	2.03	4387	12.09	11017	11.15	22764	14.86	19425	20.02	37526
Corea	3.36	876	1.28	2823	6.82	2840	3.22	8651	13.55	10010	8.48	23749
Taiwán	3.62	922	1.17	2530	7.35	3669	4.13	10625	13.83	11590	11.06	27659
Promedio	5.996	3908.2	4.642	9900.1	8.145	8665.2	12.056	22476	11.914	12330.9	17.64	30012
Desv. Est.	3.3629	2454.4	2.9531	6178.6	3.3218	4357	6.1243	9999	3.1492	5742.37	8.752	11922

Tabla 12. Conjunto completo de D3F, de acuerdo a la ecuación (3.1), al cual se le aplica el algoritmo de clasificación borrosa establecido en la sección 3.2, para definir el patrón de evolución económica de 20 países miembros de la OCDE.

CLASIFICACION BORROSA DE PARETO PARA D3F

PAIS	Objeto	DATOS 1950 (i=1)			DATOS 1973 (i=2)			DATOS 1992 (i=3)		
		U1i	U2i	U3i	U1i	U2i	U3i	U1i	U2i	U3i
Bélgica	1	0.99956	0.00026	0.00017	0.99415	0.00536	0.00049	0.02268	0.97717	0.00015
Francia	2	0.99933	0.00025	0.00041	0.98928	0.01025	0.00047	0.0005	0.99949	5.6E-06
Alemania	3	0.97339	0.00485	0.02176	0.9965	0.00325	0.00026	0.80376	0.19328	0.00296
Italia	4	0.30843	0.00771	0.68386	0.96483	0.00915	0.02601	0.98936	0.00907	0.00157
Países B.	5	0.99905	0.00068	0.00028	0.98745	0.01188	0.00067	0.89967	0.09924	0.00109
Suecia	6	0.99139	0.00721	0.0014	0.9978	0.00198	0.00022	0.98663	0.01299	0.00038
Reino U.	7	0.92893	0.06536	0.00571	0.99764	0.00192	0.00043	0.99357	0.00623	0.0002
Portugal	8	0.00034	3.4E-05	0.99963	0.0392	0.00302	0.95778	0.064	0.00761	0.92838
España	9	0.00394	0.00022	0.99584	0.39533	0.01545	0.58922	0.96324	0.0227	0.01406
E.U.	10	0.03003	0.96717	0.0028	1.4E-07	1	4.3E-09	0.01281	0.98679	0.0004
Argentina	11	0.90425	0.02634	0.06941	0.27308	0.01092	0.71601	0.0058	0.00096	0.99324
Brasil	12	0.00042	5E-05	0.99953	0.00059	8.7E-05	0.99932	0.0053	0.00152	0.99319
Chile	13	0.65533	0.0123	0.33236	0.02289	0.00176	0.97535	0.00437	0.00079	0.99485
Colombia	14	0.00071	6.8E-05	0.99922	4.5E-05	6.1E-06	0.99995	0.00017	4.2E-05	0.99979
México	15	0.00142	0.00014	0.99845	0.00017	1.9E-05	0.99981	0.0002	4.9E-05	0.99975
Venezuela	16	0.09317	0.88985	0.01697	0.80311	0.11035	0.08654	0.3531	0.03227	0.61462
India	17	0.00622	0.00101	0.99277	0.01736	0.00398	0.97865	0.02676	0.00982	0.96341
Japón	18	0.18147	0.00873	0.8098	0.93924	0.02513	0.03563	0.81077	0.18574	0.00349
Corea	19	0.00083	9.6E-05	0.99908	0.00249	0.0004	0.99711	0.18314	0.03493	0.78193
Taiwán	20	0.00107	0.00012	0.99881	0.00178	0.00025	0.99797	0.52134	0.07252	0.40613

Tabla 13. Niveles de pertenencia de la clasificación eficiente de los D3F mostrados en la Tabla 12 para cada uno de los tres años: 1950, 1973 y 1992, de 20 países miembros de la OCDE.

TESIS CON
 FALLA DE ORIGEN

DISTANCIAS A CENTROIDES DE PARETO D3F

PAIS	Objeto	DATOS 1950 (t=1)			DATOS 1973 (t=2)			DATOS 1992 (t=3)		
		(d1i) ²	(d2i) ²	(d3i) ²	(d1i) ²	(d2i) ²	(d3i) ²	(d1i) ²	(d2i) ²	(d3i) ²
Bélgica	1	0.1353	7.65744	9.61448	0.28173	3.83503	12.7071	1.39288	0.21221	17.2899
Francia	2	0.15177	9.74477	7.74924	0.30752	3.02053	14.1501	2.05484	0.04618	19.4725
Alemania	3	0.95744	13.7809	6.61481	0.20948	3.66972	13.021	0.89343	1.82193	14.733
Italia	4	2.42503	15.5977	1.71463	0.99193	10.1842	6.04081	0.35168	3.67362	8.82631
Paises B.	5	0.16076	5.82819	9.27561	0.36761	3.35188	14.1091	0.45408	1.36723	13.0266
Suecia	6	0.4862	5.33354	12.2683	0.1865	4.1897	12.5177	0.20406	1.77846	10.4477
Reino U.	7	1.28017	4.65526	15.7968	0.22499	5.12208	10.8004	0.14257	1.80052	10.0997
Portugal	8	7.91304	25.044	0.14726	6.26449	22.5881	1.26737	4.57484	13.2632	1.20117
España	9	5.33361	22.9528	0.35982	3.2315	16.3459	2.64693	0.75317	4.90639	6.23347
E.U.	10	11.5472	2.14134	37.5045	5.31939	0.00201	30.6191	4.52038	0.51498	25.5831
Argentina	11	1.42143	8.37129	5.31451	3.35518	16.7813	2.07205	6.0529	14.8615	0.46243
Brasil	12	9.43145	27.5117	0.18758	13.5484	35.3666	0.32915	16.2083	30.2981	1.18361
Chile	13	1.7552	13.0653	2.59315	6.00767	21.6758	0.92039	6.58869	15.5202	0.43657
Colombia	14	7.37414	23.8882	0.20172	12.2242	33.0958	0.08171	11.7057	23.3984	0.15254
México	15	7.17377	23.2394	0.27825	9.50708	28.4817	0.1255	11.5518	23.4883	0.16507
Venezuela	16	8.67883	2.64893	20.3597	4.23343	11.4207	12.8965	3.04138	10.0599	2.30525
India	17	15.531	38.5905	1.21031	25.8003	53.8539	3.43647	26.2077	43.2563	4.36815
Japón	18	5.77703	26.5892	2.81249	1.51016	9.23178	7.75376	0.8212	1.71571	12.5208
Corea	19	11.1894	32.971	0.31355	14.4793	36.1137	0.72395	5.22681	11.9678	2.52953
Taiwán	20	11.1536	33.26	0.35774	12.0071	32.073	0.50758	3.23073	8.66211	3.6604

Tabla 14. Distancias euclidianas a los centroides de las clases borrosas asociadas a los valores de pertenencia de los veinte Países miembros de la OCDE incluidos en el estudio, de acuerdo a la Tabla 13.

CLASE 1	CLASE 2	CLASE 3
Bélgica	Estados Unidos	Italia
Francia	Venezuela	Portugal
Alemania		España
Países Bajos		Brasil
Suecia		Colombia
Reino Unido		México
Argentina		India
Chile		Japón
		Corea
		Taiwán

Tabla 15. Clases eficientes obtenidas con el algoritmo de clasificación borrosa para los datos correspondientes a 1950, de acuerdo a los datos reportados en la Tabla 13.

CLASE 1	CLASE 2	CLASE 3
Bélgica	Estados Unidos	Portugal
Francia		España
Alemania		Argentina
Italia		Brasil
Países Bajos		Chile
Suecia		Colombia
Reino Unido		México
Venezuela		India
Japón		Corea
		Taiwán

Tabla 16. Clases eficientes obtenidas con el algoritmo de clasificación borrosa para los datos correspondientes a 1973, de acuerdo a los datos reportados en la Tabla 13.

CLASE 1	CLASE 2	CLASE 3
Alemania	Bélgica	Portugal
Italia	Francia	Argentina
Países Bajos	Estados Unidos	Brasil
Suecia		Chile
Reino Unido		Colombia
España		México
Japón		Venezuela
Taiwán		India
		Corea

Tabla 17. Clases eficientes obtenidas con el algoritmo de clasificación borrosa para los datos correspondientes a 1992, de acuerdo a los datos reportados en la Tabla 13.

CLASE 1	CLASE 2	CLASE 3
Bélgica	Estados Unidos	Portugal
Francia		España
Alemania		Argentina
Italia		Brasil
Países Bajos		Chile
Suecia		Colombia
Reino Unido		México
Venezuela		India
Japón		Corea
		Taiwán

Tabla 18. Clases eficientes obtenidas con el algoritmo de clasificación borrosa para los D3F globales (1950-1973-1992) Los países en negritas tienen el máximo nivel de pertenencia en su respectiva clase. Los niveles de pertenencia global están en la Tabla 19.

TESIS CON
FALLA DE ORIGEN

PAIS	Objeto	MATRIZ BORROSA		
		U11	U21	U31
Bélgica	1	0.83311	0.12882	0.03807
Francia	2	0.79554	0.15612	0.04834
Alemania	3	0.85701	0.09162	0.05138
Italia	4	0.73789	0.09441	0.1677
Países B.	5	0.89275	0.08316	0.02409
Suecia	6	0.90706	0.07037	0.02257
Reino U.	7	0.84231	0.11987	0.03782
Portugal	8	0.11797	0.03633	0.8457
España	9	0.4506	0.09499	0.45441
E.U.	10	0.10783	0.86755	0.02461
Argentina	11	0.37731	0.10211	0.52058
Brasil	12	0.04087	0.01719	0.94194
Chile	13	0.2033	0.05805	0.73865
Colombia	14	0.01366	0.00532	0.98102
México	15	0.0196	0.00736	0.97304
Venezuela	16	0.47398	0.31338	0.21264
India	17	0.11124	0.05536	0.8334
Japón	18	0.63807	0.13783	0.2241
Corea	19	0.09958	0.03796	0.86247
Taiwán	20	0.13912	0.04962	0.81126

Tabla 19. Valores de pertenencia borrosa obtenidos al aplicar el algoritmo de clasificación para D3F a 20 países miembros de la OCDE. Los valores en negritas indican que el país indicado tiene el nivel de pertenencia máximo en su respectiva clase.

Es decir, el modelo global de clasificación permite detectar la ocurrencia de patrones subyacentes a toda la estructura de D3F como la empleada en el presente estudio, los cuales no pudieron detectarse al considerar a los D3F como si fueran una mera colección de datos de dos formas: es decir, inconexos y temporalmente aislados.

Análisis gráfico de la solución eficiente alcanzada

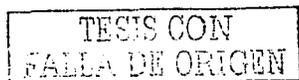
Una manera adecuada de incorporar herramientas gráficas para el análisis de los resultados obtenidos al aplicar el modelo de clasificación borrosa propuesto en este trabajo, consiste en calcular los centroides para cada una de las clases borrosas identificadas; esto se logra aplicando el Teorema 3. Una vez determinados los centroides asociados a las clases dadas en la Tabla 19, podemos estudiar la relación concomitante existente entre un par de variables o atributos de interés proyectando los centroides sobre los ejes donde se define cada una de las variables elegidas. Por ejemplo, podemos escoger como pareja de variables de interés a las empleadas en el estudio a escala reducida de la sección 3.3:

$X^{(1)}_{11}$: Años de educación formal en la población de 15 a 64 años

$X^{(1)}_{12}$: PIB per cápita

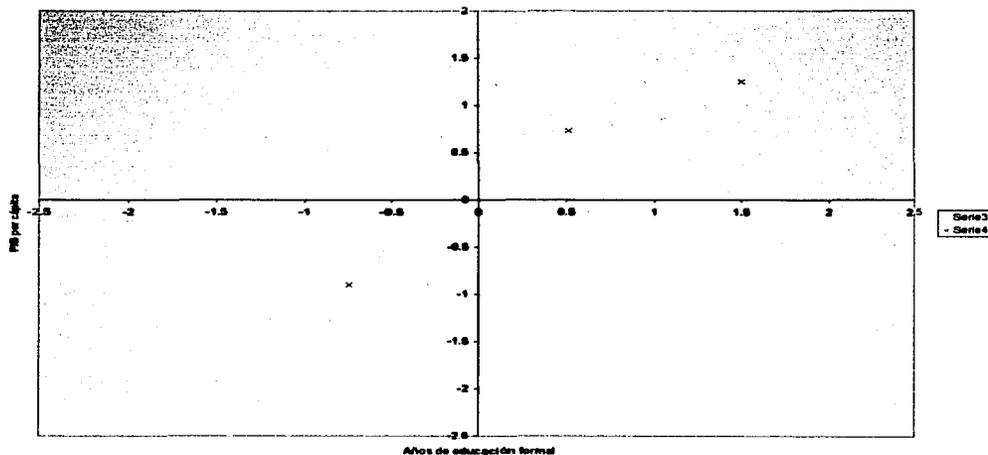
De esta manera es factible realizar comparaciones entre los resultados obtenidos en esta sección y los de la sección anterior. En la Tabla 20 se muestran los centroides temporales, obtenidos a partir de las distancias euclidianas dadas en la Tabla 14.

En la Gráfica 8 se muestran los centroides globales de las tres clases proyectados sobre las variables $X^{(1)}_{11}$ y $X^{(1)}_{12}$ conjuntamente con los valores estandarizados de ambas variables.



ESTA TESIS NO PUEDE
DE LA BIBLIOTECA

Gráfica 8. D3F (1992) proyectados sobre X1 y X2, y los centroides globales

Gráfica 8. Centroides globales (Δ) respecto a los D3F estandarizados (Δ) correspondientes a 1992 proyectados sobre los ejes de las variables X_1 y X_2 .

CLASE	AÑO	X1	X2	X3	X4
1	1950	0.83517	0.63188	0.48933	0.46309
	1973	0.65606	0.77197	0.72142	0.67534
	1992	0.34428	0.65615	0.638	0.58546
GLOBAL		0.51537	0.73044	0.83345	0.72199
2	1950	0.33567	1.90387	2.14478	2.23073
	1973	1.8981	1.80442	1.84982	1.7999
	1992	1.40329	1.14079	1.29015	1.30212
GLOBAL		1.50145	1.24971	1.07181	1.13166
3	1950	-0.7370	-0.8839	-0.8234	-0.8264
	1973	-0.7814	-0.9849	-0.9458	-0.928
	1992	-0.8692	-1.0055	-0.9749	-1.0069
GLOBAL		-0.745	-0.902	-0.9684	-0.9357

Tabla 20. Centroides temporales y globales para el conjunto D3F dados en la Tabla 12.

3.5 Reconocimiento del patrón de evolución económica de México

En esta sección final proponemos un ejercicio de reconocimiento del patrón de evolución económica de México suponiendo (hipotéticamente) que nuestro país hubiera tomado, a partir de 1951, la determinación de incrementar sustancialmente los recursos dedicados a la educación de su población, es decir, a aumentar los niveles de la variable

$X^{(1)}_{i1}$: Años de educación formal en la población de 15 a 64 años.

La idea es determinar, en este caso, la evolución económica de México identificando la clase a la cual pertenecería México al llegar al año 1992.

Supondremos que a esta variable México hubiera asignado, en 1973 y 1992, los mismos niveles que asignó Japón en dichos años:

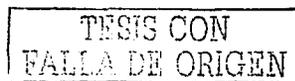
$$X^{(1)}_{i1} = 2.6, X^{(2)}_{i1} = 12.09, X^{(3)}_{i1} = 14.86$$

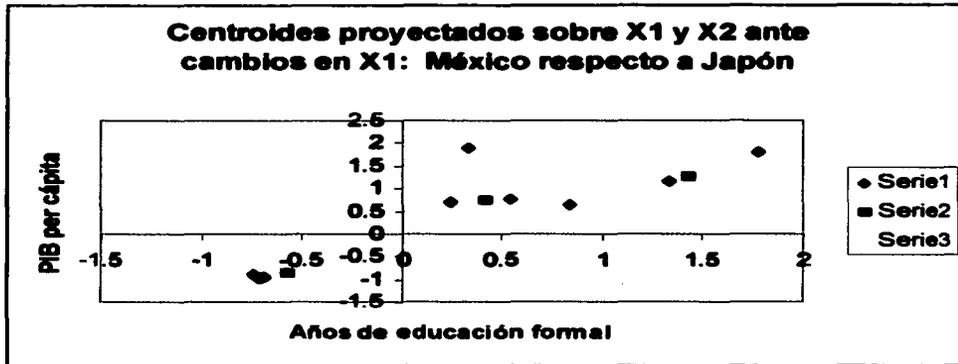
Entonces, aplicando el algoritmo de clasificación borrosa para D3F, obtendríamos los resultados mostrados en la Tabla 21.

CLASE	AÑO	X1	X2	X3	X4
1	1950	0.83517	0.63188	0.48933	0.46309
	1973	0.54771	0.76783	0.71962	0.67362
	1992	0.24292	0.69646	0.69328	0.61708
GLOBAL		0.42272	0.73547	0.84282	0.72718
2	1950	0.33567	1.90387	2.14476	2.23073
	1973	1.77509	1.80307	1.84876	1.79896
	1992	1.33295	1.14305	1.29288	1.30547
GLOBAL		1.42531	1.24903	1.06965	1.12912
3	1950	-0.737	-0.8839	-0.8234	-0.8264
	1973	-0.7024	-0.9809	-0.9517	-0.9343
	1992	-0.6889	-0.9564	-0.952	-0.9566
GLOBAL		-0.5718	-0.8687	-0.9519	-0.9053

Tabla 21. Centroides temporales y globales obtenidos al modificar para México la variable X_{i1} al mismo nivel asignado en Japón para 1973 y 1992.

En la Gráfica 9 se observan los centroides temporales y globales que se obtienen ante la modificación introducida en el vector de parámetros de México respecto a Japón.





Gráfica 9. Centroides temporales (◆) y globales (■) obtenidos al igualar la componente en X1 de México con el valor asignado a Japón para 1973 y 1992. Se muestran también los valores estandarizados (▲) modificados correspondientes de México.

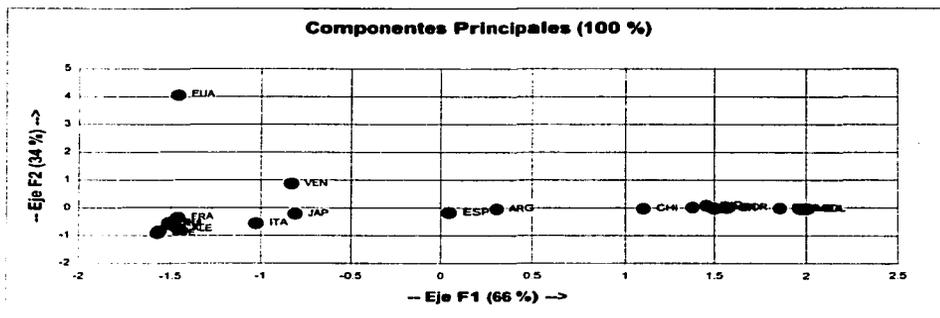
Como puede verificarse visualmente, los cambios introducidos en el vector característico de México le han permitido evolucionar económicamente en una forma totalmente distinta: se ha alejado de la clase 3 y puede ahora evolucionar hacia la clase 2 o hacia la clase 1. Es claro que el posible patrón de evolución estará fuertemente influido por las apropiadas modificaciones en el resto de las variables relevantes en el presente estudio. Quizá una adecuada política tendiente a controlar la rapidez de crecimiento poblacional permitiría al país distribuir mejor la riqueza generada, la cual se mide a través de la variable X2: PIB per cápita, influyendo así, considerablemente, en las variables X3 y X4, pues éstas varían de manera concomitante con X2.

Otra alternativa visual para presentar la información mostrada en la tabla 19, donde se incluyen los valores de pertenencia borrosa obtenidos para los D3F globales, consiste en aplicar a dichos datos el Análisis de Componentes Principales (ACP) [Kaufman-Rousseeuw, 1990]. En este caso, el número de componentes principales no degenerados es igual al número de clases o conglomerados menos uno, i.e. (K-1); debido a que, para cada objeto, la suma de pertenencias es una constante. Consecuentemente, obtendremos en nuestro estudio dos componentes principales. Al aplicar el ACP al de datos dados en la tabla 19, empleando un programa estándar como el XLSTAT™ (Versión 6.0), se obtienen los componentes principales que se muestran en la tabla 22 y que se representan en la gráfica 10. En esta gráfica 10 se confirma que:

1. E.U.A. es el único integrante de una sola clase
2. España y Argentina están en la frontera entre las dos clases restantes
3. México se encuentra, junto con Colombia, en el extremo de la tercera clase.

Componentes principales de μ_{ki}

	F1	F2
BEL	-1.503	-0.523
FRA	-1.462	-0.340
ALE	-1.458	-0.737
ITA	-1.024	-0.564
PB	-1.561	-0.826
SUE	-1.569	-0.908
RU	-1.505	-0.579
POR	1.501	-0.019
ESP	0.047	-0.177
EUA	-1.462	4.060
ARG	0.295	-0.044
BRA	1.859	-0.010
CHI	1.104	-0.027
COL	2.003	-0.031
MEX	1.973	-0.029
VEN	-0.829	0.860
IND	1.458	0.083
JAP	-0.807	-0.218
COR	1.564	0.014
TAI	1.374	0.018

Tabla 22. Componentes principales de los valores μ_{ki} mostrados en la Tabla 19.

Gráfica 10. Gráfica de los componentes principales dados en la Tabla 22.

TESIS CON
FALLA DE ORIGEN

De hecho, si bien la proyección de los valores graficados sobre F2 nos permite verificar la constitución de E.U.A. en una sola clase, no es claro cómo esto puede ser explicado definiendo a F2 en términos de las cuatro variables empleadas en el presente trabajo.

Aunque en este trabajo se consideren sólo cuatro variables o atributos, el modelo de clasificación propuesto puede admitir un número mayor de éstas, la única limitación al respecto sería por la disponibilidad de equipo computacional apropiado.

En el ejemplo hipotético de esta sección hemos comprobado cómo el modelo de clasificación borrosa propuesto permite, no sólo describir la evolución en el tiempo de un conjunto de D3F, sino además emplear la metodología conocida como reconocimiento de patrones. Es decir, permite clasificar objetos respecto a un conjunto de atributos o variables que varían en el tiempo, y, posteriormente, podemos utilizar la clasificación desarrollada para identificar ya sea nuevos objetos respecto a la clasificación encontrada, o modificar los vectores de algún objeto ya considerado para efectuar investigaciones de escenarios posibles. En otras palabras podemos considerar escenarios del tipo "¿Qué pasa si ...?", como el que presentamos al modificar el vector característico de México.

Conclusiones

Hemos propuesto un modelo de clasificación borrosa que permite analizar el problema de clasificar información presentada en forma de D3F. La estructura algorítmica del modelo permite una secuenciación que facilita el control del procedimiento iterativo asociado.

Como se observa en (2.22)-(2.23), el algoritmo propuesto sólo depende de los parámetros (μ_{ki}) y $d(x_i, v_k)$, esta última es una distancia euclidiana que puede interpretarse como un coeficiente de disimilitud y que no requiere que los atributos sean variables medibles a escala de intervalo, i.e. que sean variables continuas, por lo tanto pueden incluirse en estudios de este tipo variables binarias [Kaufman-Rousseeuw, 1990]. Lo anterior significa que, si en particular se incluyeran variables binarias, éstas pueden ser consideradas como variables continuas aplicando las métricas apropiadas: euclidiana, Manhattan, o incluso euclidiana ponderada. Sin embargo, para estos casos existen algoritmos de clasificación diseñados específicamente para datos binarios de dos formas (D2F), sobre todo si las variables o atributos restantes deben ser considerados como rígidos o no borrosos, por ejemplo el análisis de tipo jerárquico llamado monotético [Kaufman-Rousseeuw, 1990]. No obstante, en tal caso debe considerarse que, al incluir variables binarias, éstas pueden ser simétricas (aquellas cuyos posibles resultados tienen asignado el mismo nivel de ponderación) o asimétricas (en las cuales se tienen asignados distintos niveles de ponderación para cada resultado). Por lo tanto, al considerar a cada variable binaria como si fuera continua, se asume que cada una de ellas es simétrica.

Por otra parte, una de las preguntas que surgen en problemas de clasificación como los tratados en este trabajo es acerca de cómo medir la influencia que tienen los n atributos o variables empleadas en la clasificación de los m objetos. Al respecto podemos ponderar esta influencia a través del ACP del conjunto de valores de pertenencia (μ_{ki}) obtenidos. El ACP nos permite resumir la gran cantidad de información obtenida y, posteriormente, visualizarla en una gráfica de dimensión $(K-1)$. Por ejemplo, si $K=3$ (como en el presente caso), se obtiene una gráfica bidimensional, de los dos componentes principales no

TESIS CON
FALLA DE ORIGEN

degenerados para la matriz $U=(\mu_{ki})$ de la tabla 19, que se muestra en la gráfica 10. Si se tienen más de tres clases o conglomerados, $K>3$, se tendrán más de dos componentes principales y podrían entonces graficarse las dos componentes con los mayores valores propios; de esta manera se podría explicar la máxima cantidad posible de variabilidad de los valores (μ_{ki}) , que es el objetivo central del ACP [Carroll et al, 1997]. Este enfoque nos motivaría entonces para llegar a una interpretación de las dos componentes principales en términos de nuevas variables que pueden definirse dentro del contexto en el cual se han definido las variables originalmente incluidas en el estudio.

Es conveniente enfatizar que el procedimiento algorítmico de actualización propuesto debe continuarse hasta que los cambios observados entre dos iteraciones sucesivas sean más pequeños que algún valor umbral previamente establecido. Al respecto, han sido reportados los resultados de convergencia del algoritmo establecido en el teorema 2 [Gordon, 1999]. Adicionalmente, con el fin de identificar valores apropiados de K , se han propuesto diversas medidas, tanto teóricas como de simulación, de la borrosidad de una solución [Pal-Bezdek, 1995]; aunque, obviamente, estos estudios no han sido realizados para el algoritmo propuesto en este trabajo en el teorema 3.

Finalmente, respecto al modelo algorítmico propuesto para D3F en el presente trabajo, es suficientemente flexible como para adaptarse a los recursos computacionales disponibles y, al mismo tiempo, permite contrastar los resultados obtenidos con aquellos correspondientes a escenarios diversos al contexto original en el que surgen los datos. Esto permite, por ejemplo, considerar nuevas situaciones T , o bien emplear otros valores en los parámetros del modelo, como el nivel de borrosidad empleado, para mejorar así la bondad de la clasificación obtenida.

Respecto a los objetivos específicos planteados en la sección 3.1 para este estudio a escala completa de los D3F considerados, podemos concluir que:

1. Se ha logrado una descripción global de la economía global válida no sólo durante el período histórico 1950-1992, sino también válida para períodos más amplios; al permitir considerar no sólo variables estrictamente económicas, sino además variables socio-económicas que, por su propia naturaleza, son intrínsecamente borrosas.
2. El modelo matemático propuesto permite identificar las tendencias de convergencia y divergencia, en los niveles de las cuatro variables incluidas, de las economías de los países considerados en el presente estudio. Por ejemplo, México muestra un comportamiento sistemático (prácticamente invariable) hacia la clase 3, con una tendencia a alejarse de la clase 1, al menos hasta 1992; en cambio, E.U. ha llegado a constituir *per se* una sola clase, la clase 2, en la cual es el líder absoluto con un nivel de pertenencia de .86755.
3. Consideramos que el presente estudio aporta elementos de juicio para establecer en el futuro las características de las diversas fases del desarrollo económico de los países, específicamente de aquellos integrantes de la OCDE.

De esta manera, el hecho de que el modelo de clasificación propuesto incorpore la borrosidad de la pertenencia de los objetos respecto a las distintas clases identificadas es sumamente útil, pues reproduce la incertidumbre natural que poseen los objetos respecto a los atributos, ya sean borrosos o no, considerados en el estudio. Después de todo, esperamos que la mayoría de los fenómenos que nos interesa conocer posean, en mayor o menor medida, una cierta estructura subyacente que esperamos descubrir con ayuda de herramientas como la aquí presentada.

TESIS CON
FALLA DE ORIGEN

Bibliografía

1. Bezdek, J.N. (1981), *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York
2. Bowman, V. J. (1976), "On the relationship of the Tchebycheff norm and the efficient frontier of multiple-criteria", in *Multiple Criteria Decision Making*", pp. 78-86, Thieriez H. and Zionts S. Editors, Springer-Verlag, Berlin
3. Carroll, J.D. and Arabie, P. (1983), "INDCLUS: An individual differences generalization of the ADCLUS model and MAPCLUS algorithm", *Psychometrika*, Vol. 48, pp. 157—169
4. Carroll, J.D., Green P.E. and Chaturvedi A. (1997), *Mathematical tools for applied multivariate analysis*, Academic Press, San Diego
5. Da Cunha N.C. and Polak, E. (1967), "Constrained minimization under vector valued criteria in finite-dimensional space", *J. Math. Anal. Appl.*, Vol. 19, pp.103-124
6. Dunn, J.C. (1973), "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters", *J. Cybernetics*, Vol. 3, pp.32-57
7. Dunn, J.C. (1974), "Well separated clusters and optimal fuzzy partitions", *J. Cybernetics*, Vol. 4, pp. 85-104
8. Gordon, A.D. (1999), *Classification*, 2nd Edition, Chapman and Hall, London
9. Kaufman, L. and Rousseeuw, P.J. (1990), *Finding Groups in Data*, Wiley, New York
10. Looney, C. (1997), *Pattern Recognition Using Neural Networks*, Oxford Univ. Press, New York
11. Maddison, A. (1997), *La Economía Mundial (1820-1992): Análisis y Estadísticas*, Centro de Desarrollo, O.C.D.E., París, Francia
12. Pal, N.R. and Bezdek, J.C. (1995), "On the cluster validity for the fuzzy c-means model", *IEEE Transactions on fuzzy systems*, Vol. 3, 370-379.
13. Rockafeller, R.T. (1970), *Convex Analysis*, Princeton University Press, Princeton, N.J.
14. Ross, T.J. (1995), *Fuzzy Logic with Engineering Applications*, Mc Graw-Hill, N.Y.
15. Ruspini, E.H. (1969), "A new approach to clustering", *Inform. Control*, Vol.15, pp. 22-32
16. Shapiro, J.F. (1979), *Mathematical Programming: Structures and Algorithms*, Wiley, New York
17. Tucker, L.R. (1972), "Relation between multidimensional scaling and threemode factor analysis", *Psychometrika*, Vol. 37, pp.3-27
18. Yu, P.L. (1973), "A class of solutions for group decision problems", *Management Science*, Vol. 19, pp. 936-946
19. Zadeh, L. (1965) "Fuzzy sets", *Information and Control*, Vol. 8, pp.338-353

TESIS CON
 FALLA DE ORIGEN