

Aplicación del Análisis de Escalamiento Multidimensional

Autor: Lic. Juan Carlos Flores López

Para el desarrollo de esta investigación se tomo en cuenta el estudio realizado en la investigación sobre las causas del abandono de las materias de los estudiantes de la Facultad de Ciencias Puras y Naturales, realizado para la gestión anterior.

La teoría del análisis de escalamiento multidimensional, demanda el conocimiento de distancias y las medidas de similaridad, por ejemplo, la distancia euclidea, distancia euclidea al cuadrado, Distancia de Minkowski, Distancia city block o «Manhattan», Medidas de similaridad para datos binarios, etc.

El análisis de escalamiento multidimensional (MDS)-*multidimensional scaling*- es una técnica de reducción de datos como otras como es el análisis factorial o análisis de componentes principales, por ejemplo. El objetivo principal del MDS es representar N objetos en un espacio dimensional reducido (q dimensiones, siendo $q < N$), de tal forma que la distorsión causada por la reducción de la dimensionalidad sea la menor posible, es decir, que las distancias entre los objetos representados en el espacio q dimensional, sean lo más parecidas posible a las distancias en el espacio N dimensional.

Dado que será difícil que las distancias coincidan, el objetivo del MDS es conseguir que ambas configuraciones dimensionales sean lo más parecidas posible. Para ello será necesario construir un indicador de esa proximidad que se denomina *stress* y otro similar el *s-stress*,

cuya interpretación de resultados se presenta en el siguiente cuadro.

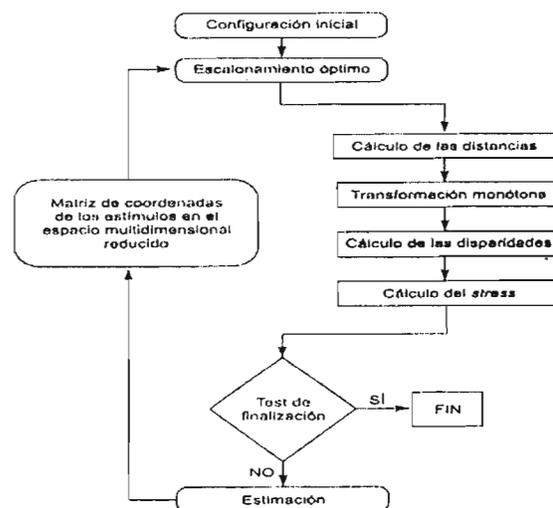
Cuadro. Interpretación de Stress en términos de bondad de ajuste del MDS.

Stress	Bondad de ajuste
0,2	Malo
0,1	Mínimo Razonable
0,05	Bueno
0,025	Excelente
0	Perfecto

El indicador *S-Stress*, esta dado por:

$$S - Stress = \sqrt{\frac{\sum_{i \neq j} (d^2_{ij} - \delta^2_{ij})^2}{\sum_{i \neq j} d^4_{ij}}}$$

A continuación se muestra el Síntesis del algoritmo básico del MDS, para el calculo del indicador.



1. Estimación del modelo

En el caso en que no se haya concluido el proceso iterativo, en esta segunda fase se procede a la estimación del modelo. Con el algoritmo ALSCAL, como todos los diseñados para resolver el problema que se plantea el MDS es iterativo, debiéndose determinar en cada iteración dos variables fundamentales: las *coordenadas de los estímulos* y los *pesos de los individuos*, en el caso de que se trate de un *escalonamiento multidimensional ponderado*, los pesos y las coordenadas no pueden obtenerse simultáneamente. Por ello la estimación del modelo se hace en dos etapas: (a) estimación de los pesos, y (b) estimación de las coordenadas de los estímulos. En un modelo no ponderado, lógicamente, no se aplica la primera fase. (el proceso teórico puede ser visto en la Investigación de la gestión 2009 de título Aplicación del Método Escalamiento Multidimensional).

2. Aplicación del análisis de Escalamiento Multidimensional

La teoría del análisis de Escalamiento Multidimensional, se toma en cuenta en el estudio realizado sobre la investigación de las causas del abandono de las materias de los estudiantes de la Facultad de Ciencias Puras y Naturales.

En la actualidad la Facultad de Ciencias Puras y Naturales (FCPN), es parte de las 13 Facultades que tiene la Universidad Mayor de San Andrés. La FCPN tiene 6 Carreras que son las siguientes: ESTADISTICA, BIOLOGIA, INFORMATICA, FISICA, MATEMATICA y QUIMICA

De acuerdo a estudios realizados, el abandono ocurre por diversos motivos, estos pueden deberse a entorno

universitario, como el entorno social, cultural y familiar en el que se mueve el alumno. Según el informe elaborado por la Comisión Europea, los factores que provocan el abandono de los estudios temprano se pueden clasificar del siguiente modo: Características individuales propias de los alumnos, Razones educativas, Razones familiares, Comunidad y amigos.

La deserción universitaria es un problema educativo que afecta al desarrollo de la sociedad, y se da principalmente por falta de recursos económicos y por una desintegración familiar.

Es un fenómeno social ocasionado por diversas causas ya sean políticas, económicas, familiares, etc. Lo cual debe ser estudiado detenidamente para determinar las posibles soluciones, así como también su prevención. El abandono temporal o definitivo de los estudios de un individuo es motivo de varios factores, además de otros elementos ya sean externos o internos que motivan a tomar esta decisión. Los factores externos, por ejemplo pueden ser por presiones económicas, influencia negativa de padres, amigos, familiares, embarazos a temprana edad, maestros, complejidad de las materias, etc.

Y los factores internos pueden ser por desinterés personal, no tener motivación en la vida, desagrado por la universidad, carrera, materia (desgano de salir adelante y sobresalir entre las demás personas, cruce de horarios, etc.).

Para que un estudiante abandone una carrera universitaria se combinan aspectos como el lugar en donde reside, el nivel de ingresos, el nivel educativo de los padres de familia, la necesidad de trabajar para mantenerse o contribuir a los ingresos familiares y el propio ambiente familiar,

incluso de violencia en el que se vive. "Esta situación es la que afecta con mayor fuerza a los jóvenes de menores ingresos, por lo que el tema financiero y la eficiencia en el gasto se hace más crítico".

La "deficiente preparación previa (en el bachillerato)" es otra de las causas del abandono de los estudios universitario; la carencia de mecanismos de financiamiento o becas estudiantiles; la prevalencia de políticas de "ingreso irrestricto, selectivo sin cupo fijo o selectivo con cupo"; el desconocimiento de lo que es la profesión, el ambiente escolar y la carencia de lazos afectivos con la universidad, también impactan en los jóvenes para que dejen la Universidad.

3. Variables en consideración de la investigación

En esta investigación se ha considerado la encuesta sobre el abandono de materias de los estudiantes de la Facultad de Ciencias Puras y Naturales de la Universidad de Mayor de San Andrés. La muestra fue de tamaño 269 Estudiantes. Se realizó un muestreo aleatorio simple considerando dominios de grupo, además dentro de cada dominio se distribuye la muestra de manera proporcional a los tamaños de los dominios. Considerando que la carrera de Informática es muy numerosa, se ha visto por conveniente tomar la mitad de la muestra en informática y la otra mitad se distribuyó proporcionalmente al tamaño de cada carrera, con el objeto de tener una muestra más representativa. El resultado del muestreo es el siguiente Carrera de Estadística 22 estudiantes, Biología 32 estudiantes, Informática 134 estudiantes seleccionados, Física 16 estudiantes seleccionados, Matemáticas 43 estudiantes seleccionados, Química 22 estudiantes seleccionados.

Las variables más importantes, que se toma en cuenta en este estudio son:

¿asiste regularmente a clases?, ¿considera cómodos sus horarios de clases?, ¿considera cómodos sus horarios de clases?, ¿dispone de tiempo suficiente para realizar trabajos y/o prácticas universitarias?, ¿alguna vez abandonó una o varias materias?.

Si usted considerara abandonar una o varias materias, ¿qué lo motivaría a abandonarlas?

- 1) Falta de tiempo
- 2) Flojera
- 3) Bajo rendimiento
- 4) Falta de interés
- 5) Por trabajo
- 6) Docente
- 7) Choque de horario
- 8) Desorganizado
- 9) Otros _____

Si usted considerara abandonar sus estudios universitarios, ¿por qué lo haría?

- 1) Falta de tiempo
- 2) Falta de dinero
- 3) Bajo rendimiento
- 4) Otros _____

Para el análisis, se ha pedido a los estudiantes, que nos valoren el por qué un estudiante abandona una o más materias bajo la consideración de los siguientes criterios, atendiendo a la similitud con que las perciben. Para ello utiliza una escala de 1 (totalmente diferentes) a 7 (idénticas).

La siguiente matriz de *disparidades originales* - o *proximidades* - nos muestra las medias de las puntuaciones ofrecidas por los estudiantes de la muestra de la Facultad de Ciencias Puras y Naturales.

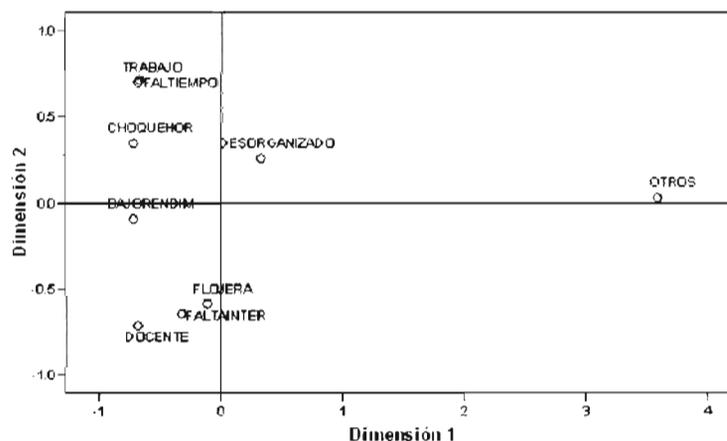
	falta de tiempo	Flojera	Bajo rendimiento	Falta de interés	Por trabajo	Docente	Choque de horario	Desorganizado	otros
falta de tiempo	7	3,26	3,85	3,19	4,56	3,2	4,07	3,51	2,38
Flojera	3,26	7	4,25	3,92	3,15	3,27	3,3	3,46	2,32
Bajo rendimiento	3,85	4,25	7	4,21	4,21	4	4	3,62	2,53
Falta de interés	3,19	3,92	4,21	7	3,19	3,91	3,33	3,38	2,38
Por trabajo	4,56	3,15	4,21	3,19	7	3,19	4,1	3,19	2,42
Docente	3,2	3,27	4	3,91	3,19	7	3,73	3,02	2,19
Choque de horario	4,07	3,3	4	3,33	4,1	3,73	7	3,46	2,3
Desorganizado	3,51	3,46	3,62	3,38	3,19	3,02	3,46	7	2,6
Otros	2,38	2,32	2,53	2,38	2,42	2,19	2,3	2,6	7

Nótese que en la diagonal de la matriz aparecen siete porque es una percepción de similitud de cada variable, siempre ha de ser idéntica a sí misma. Se puede observar que esta matriz es simétrica por construcción. Si creamos un mapa en dos dimensiones para ilustrar mejor la percepción de los estudiantes por que abandonan su(s) materias en un semestre, este mapa debería representar como puntos cercanos a las variables ¿por que abandonaría una materia un estudiante?. De

acuerdo a los resultados de la matriz las variables de abandono seria la falta de tiempo y por trabajo, porque la disparidad entre ellas es alta (4,56), y así sucesivamente.

Considerando las variables anteriormente mencionadas y utilizando el software SPSS se procede a la utilización del método de análisis de escalonamiento multidimensional, cuyos resultados son los siguientes:

Figura 1. Modelo de distancia euclidea



La Figura 1 ofrece el mapa que se obtiene al representar las coordenadas bidimensionales resultantes de aplicar a la matriz anterior uno de los algoritmos que existen para efectuar un MDS, el ALSCAL. Puede comprobarse que los estudiantes de la Facultad de Ciencias Puras y Naturales abandonan su(s) mate-

ria(s), por trabajo, por falta de tiempo y choque de horario, o por falta de interés, flojera y/o por su docente.

El indicador de calidad de método de escalamiento MDS (*stress* o *s-stress*) refleja este hecho, como podemos ver la salida del software

Iteration history for the 2 dimensional solution (in squared distances)
 Young's S-stress formula 1 is used.

Iteration	S-stress	Improvement
-----------	----------	-------------

Iterations stopped because
 S-stress improvement is less than ,001000

Una medida de las discrepancias entre las matrices de disparidades y distancias de acuerdo a la teoría la calidad de la representación lograda por el MDS, es el estadístico denominado S- Stress, que fue propuesto por Takane, Young y De Leeuw (1977), autores del algoritmo ALSCAL y que, por ello, es la función que se minimiza en este algoritmo:

El valor del *s-stress* está siempre comprendido entre 0 y 1 y cualquier valor inferior a 0,1 indica que la solución

obtenida es una buena representación de los objetos de la solución *N* dimensional inicial.

En nuestro caso el S-Stress es 0.001 lo que implica una excelente representación de los objetos, es decir la similitud de la percepción de los estudiantes por que abandonan sus materias de los estudiantes de la muestra es excelente, bajo el criterio del método en consideracion.

Stress and squared correlation (RSQ) in distances
 For matrix
 Stress = ,05644 RSQ = ,99210

Se puede ver también que el indicador de bondad de ajuste que nos proporciona el software RSQ = 0.99210. Es muy buena, es decir la solución dimensional reducida es una excelente representación de la solución *N* dimensional, si la ordenación de las distancias entre los objetos de la primera mantiene la ordenación de las disparidades originales de la segunda. En el caso ideal, el grafico de dispersión que representa distancias y disparidades debería

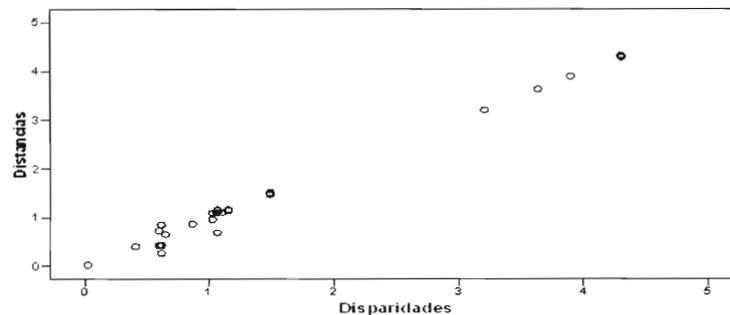
ser una línea recta como indica la teoría, lo cual se cumple en esta investigación. Los siguientes datos muestran la solución bidimensional final resultante de la aplicación del análisis de escalamiento multidimensional a los datos de nuestra investigación. En este cuadro aparecen las coordenadas de cada objeto (percepción de los estudiantes del por qué abandonan los estudiantes sus materias). A partir de estas coordenadas se derivan las matrices de distancias euclideas.

Configuration derived in 2 dimensions

Stimulus Number	Stimulus Name	Stimulus Coordinates	
		Dimension 1	Dimension 2
1	FALTIEMP	-,6918	,7502
2	FLOJERA	-,0324	-,5889
3	BAJOREND	-,7376	-,1124
4	FALTAINT	-,2837	-,6747
5	TRABAJO	-,7044	,7293
6	DOCENTE	-,6977	-,7580
7	CHOQUEHO	-,7455	,3262
8	DESORGAN	,3490	,2891
9	OTROS	3,5440	,0391

La Figura 2 muestra la dispersión que representa distancias y disparidades de nuestros datos, el cual efectivamente se ajusta a una línea recta, como indica la teoría MDS.

Figura 2 Ajuste lineal modelo de distancia euclidea



Como resultado de la investigación, utilizando el método de Escalamiento Multidimensional, se puede ver que los estudiantes de la Facultad de Ciencias Puras y Naturales abandonan sus materias por trabajo, por falta de tiempo y choque de horario, y un segundo motivo por falta de interés, flojera y/o por su docente.

Para la comprobación de los resultados se ha utilizado el indicador de calidad de método de escalamiento MDS

(*stress* o *s-stress*) propuesto por Takane, Young y De Leeuw autores del algoritmo ALSCAL que es la función que se minimiza en este algoritmo, cuyo resultado fue de 0.001 lo que implica una excelente representación de los objetos. Es decir la similitud de la percepción de los estudiantes del por qué abandonan sus materias de los estudiantes de la muestra por el método Escalamiento Multidimensional es excelente.

"Conseguimos obtener así la fórmula estadística para conocer aproximadamente la posición de un electrón en un instante determinado. Pero, personalmente, no creo que Dios juegue a los dados."

Albert Einstein