Introducción.

Bajo la denominación de Análisis de Datos Multivariantes, se recogen diversas técnicas que tienen un denominador común y es presentar los datos originales de forma resumida y a través de representaciones gráficas que faciliten su comprensión. Estamos en el lado exploratorio del Análisis Multivariante aunque más adelante comprobaremos las interacciones establecidas con el lado confirmatorio clásico.

Como es sabido el Análisis Multivariante clásico gira en torno a poblaciones multivariantes bajo tres condicionantes: Normalidad, Análisis basado en vectores media y matrices de covarianzas, y Estructura lineal con todo el desarrollo algebraico. Pero a medida que se han ido extendiendo los campos de aplicación de la Estadística, han aparecido muchos tipos de datos que no admiten esos condicionantes como hipótesis plausibles. Por otro lado los datos multivariantes se recogen porque es más rápido que observar una variable adecuada, que incluso puede no conocerse o depender de otras. Por lo tanto los datos multivariantes son recogidos para hacer un estudio preliminar de un fenómeno y a veces las conclusiones que se obtienen son una guía para elegir posteriormente un estudio multivariante convencional: un simple diagrama de dispersión entre las variables nos dirá si la relación entre ambas es lineal, antes de realizar un análisis de regresión bivariante.

Esta proliferación y desarrollo del carácter emergente del Análisis Multivariante, ha sido posible gracias a la evolución y potencia de los medios informáticos, y entre la gran variedad de técnicas podemos encontrar el análisis de componentes principales, análisis de correspondencias simple, análisis de correspondencias múltiple, análisis biplot, escalamiento multidimensional, escala dual, escala óptima, análisis conceptual formal, análisis cluster y diversas variantes de todos ellos. La gran mayoría de ellas se pueden clasificar en dos grandes grupos: los métodos factoriales o de reducción de dimensión y los métodos de clasificación. Los primeros utilizan unos cálculos de ajuste que recurren esencialmente al álgebra lineal, y producen unas representaciones gráficas donde los objetos a describir se transforman en puntos sobre un eje o sobre un plano, proporcionando unas representaciones sintéticas de amplios conjuntos de valores numéricos, y los segundos ponen en juego una formulación y unos cálculos algorítmicos que producen unas clases o unas familias de clases, que permiten agrupar y ordenar los objetos a describir.

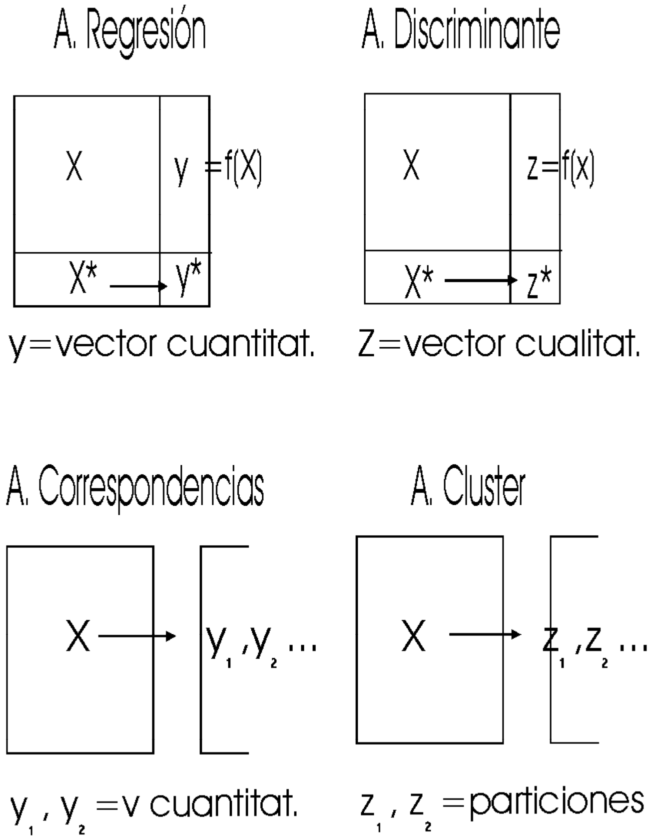
Estas dos familias de métodos son más complementarias que concurrentes, y pueden ser utilizadas de forma simultánea sobre un mismo conjunto de datos. Cada una de ellas da un punto de vista diferente sobre los materiales estadísticos que le son proporcionados.

Las técnicas factoriales tienen una parte de fundamentación general común a todas ellas, que hoy se conoce con el nombre de 'Aproximación de una matriz por otras de bajo rango´ y está basada en la teoría general de la Descomposición en Valores Singulares (DVS) de una matriz.

Para globalizar estas ideas y mostrar la relación existente entre estas técnicas con análisis canónicos bien conocidos como análisis de regresión múltiple o análisis discriminante, tenemos el diagrama siguiente:
Sea $X_{I\times J}$ una matriz de observaciones, esquemáticamente tendríamos el cuadro de la figura 1.

- El Análisis de Regresión Múltiple: estudia las relaciones entre una variable dependiente, y, vector de valores cuantitativos, y un conjunto de variables independientes X, y predice y^* de X^* dado.
- El Análisis Discriminante es igual, solo que la variable dependiente z es cualitativa, es decir intentamos clasificar un conjunto de observaciones.
- El Análisis de Correspondencias, Componentes Principales, Escalas, etc, producen conjuntos de valores cuantitativos (coordenadas) para las variables o sujetos.
- El Análisis Cluster, produce conjuntos de valores cualitativos, en el sentido de particiones de las variables o sujetos.
- Obsérvese la diferencia entre el A. Discriminante, donde hay una partición concreta conocida de antemano y el A. Cluster, donde las particiones son generadas por el análisis.
- El A. Discriminante se puede considerar la forma discreta del A. Regresión, y el A. Cluster, la forma discreta de las escalas multidimensionales (A. Correspondencias), en el sentido de que el A. Correspondencias genera conjuntos de valores en forma de puntos en el espacio multidimensional, mientras las técnicas clusters generan valores discretos que asignan los sujetos a los grupos.

Figure 1: Relación entre los diferentes análisis



Bibliografía.

- **Everit, B.S.** (1992). *The Analysis of Contingency Tables.* Chapman Hall.
- **Foucart, T.** (1982). *Analyse Factorielle.* Masson.
- **Greenacre, M.** (1984). *Theory and Applications of Correspondence Analysis.* Academic Press.
- **Greenacre, M.** (1993). *Correspondence Analysis in Practice.* Academic Press.
- **Gutiérrez-González-Torres-Gallardo** (1994). *Técnicas de Análisis de datos multivariable. Tratamiento computacional.* Universidad de Granada.
- **Hair, J.** (2000). *Análisis Multivariante.* Prentice Hall.
- **Lebart-Morineau-Fenelon, L.** (1985). *Tratamiento estadístico de Datos.* Marcombo.
- **Lefebvre, J.** (1976). *Introduction aux Analyses Statistiques Multidimensionnelles.* Masso.
- **Mardia, K.** (1979). *Multivariate Analysis.* Academic Press.
- **Peña, D.** (2002). *Análisis de datos multivariantes.* McGraw-Hill.
- **Volle, M.** (1989). *Analyse des données.* Economica.

Producto con Matrices.

1. **Producto de una matriz A por una matriz diagonal $\mathbb{D}_\lambda$.**
   - $(\mathbb{D}_\lambda)_{I\times I}(A)_{I\times J} \equiv$ Se multiplican las filas de la matriz por los elementos de la diagonal.
   - $(A)_{I\times J}(\mathbb{D}_\lambda)_{J\times J} \equiv$ Se multiplican las columnas de la matriz por los elementos de la diagonal.
2. **Producto de una matriz A por el vector $\vec{1}$.**
   - $(A)_{I\times J}\,(\vec{1})_{J\times 1} \equiv$ Suma de las filas de la matriz A, (vector columna).
   - $(A^{'})_{J\times I}(\vec{1})_{I\times 1} \equiv$ Suma de las columnas de la matriz A, (vector columna).
   - $(\vec{1}^{'})_{1\times I}(A)_{I\times J} \equiv$ Suma de las columnas de la matriz A, (vector fila).
   - $(\vec{1}^{'})_{1\times J}(A^{'})_{J\times I} \equiv$ Suma de las filas de la matriz A, (vector fila).
3. **Producto de una matriz por un vector.**
   - Operaciones que **no** se pueden realizar:
     - $(\vec{v})_{I\times 1}(A)_{I\times J}$
     - $(A)_{I\times J}(\vec{v}^{'})_{1\times I}$
   - Operaciones que se pueden realizar:
     - $(A)_{I\times J}(\vec{v})_{J\times 1} =$ vector columna. Producto de las filas de la matriz por los elementos de $\vec{v}$.
     - $(A^{'})_{J\times I}(\vec{v})_{I\times 1} =$ vector columna. Producto de las columnas de la matriz por los elementos de $\vec{v}$.
     - $(\vec{v}^{'})_{1\times I}(A)_{I\times J} =$ vector fila. Producto de las columnas de la matriz por los elementos de $\vec{v}$.
     - $(\vec{v}^{'})_{1\times J}(A^{'})_{J\times I} =$ vector fila. Producto de las filas de la matriz por los elementos de $\vec{v}$.
4. **Producto de una matriz diagonal por un vector $\vec{v}$.**
   - $(\mathbb{D}_{\vec{v}}^{-1})_{I\times I}(\vec{v})_{I\times 1} = (\vec{1})_{I\times 1} \equiv$ Se multiplica cada elemento de la diagonal por las filas del vector obteniéndose un vector columna.
   - $(\vec{v}^{'})_{1\times I}(\mathbb{D}_{\vec{v}}^{-1})_{I\times I} = (\vec{1}^{'})_{1\times I} \equiv$ Se multiplica cada elemento de la diagonal por las columnas del vector obteniéndose un vector fila.
5. **Producto de una matriz diagonal $\mathbb{D}_\lambda$ por el vector $\vec{1}$.**
   - $(\mathbb{D}_\lambda)_{I\times I}(\vec{1})_{I\times 1} = (\vec{v_\lambda})_{I\times 1} \equiv$ Vector columna cuyos elementos son los de la matriz diagonal.
   - $(\vec{1}^{'})_{1\times I}(\mathbb{D}_\lambda)_{I\times I} = (\vec{v_\lambda}^{'})_{1\times I} \equiv$ Vector fila cuyos elementos son los de la matriz diagonal.
6. **Producto de un vector $\vec{v}$ por el vector $\vec{1}$.**
   - $(\vec{1})_{I\times 1}(\vec{v}^{'})_{1\times J} \equiv$ Matriz cuyas filas son el vector $\vec{v}^{'}$.
   - $(\vec{v})_{I\times 1}(\vec{1}^{'})_{1\times J} \equiv$ Matriz cuyas columnas son el vector $\vec{v}$.
   - $(\vec{1}^{'})_{1\times I}(\vec{v})_{I\times 1} \equiv$ Escalar igual a la suma de las componentes del vector $\vec{v}$.
   - $(\vec{v}^{'})_{1\times I}(\vec{1})_{I\times 1} \equiv$ Escalar igual a la suma de las componentes del vector $\vec{v}^{'}$.
7. **Producto de vectores.**
   - $(\vec{v}^{'})_{1\times I}(\vec{v})_{I\times 1} =$ escalar.
   - $(\vec{v})_{I\times 1}(\vec{v}^{'})_{1\times I} =$ Matriz.
8. **Traza de una matriz.**
   - $\mathrm{tr}(XX^{'}) = \mathrm{tr}(X^{'}X) = \sum_{i=1}^{n} x_i^2$ (suma de los cuadrados de todos sus términos.)
   - $\mathrm{tr}(A-B) = \mathrm{tr}(A) - \mathrm{tr}(B)$
   - $\mathrm{tr}(A-B)^{'} = \mathrm{tr}(A)^{'} - \mathrm{tr}(B)^{'}$
   - $\mathrm{tr}(A\cdot B) = \mathrm{tr}(B\cdot A)$