

UN ENFOQUE GEOMÉTRICO AL ANÁLISIS DE DATOS, Profesor Cecilio Mar Molinero (UAB).

"En la actualidad, dada la gran disponibilidad de herramientas informáticas, es fácil recoger información, que se suele almacenar en bases de datos. Las empresas y administraciones públicas guardan datos sobre clientes, usuarios y personas que han demostrado un cierto interés en algún tema. Pero almacenar los datos no es suficiente, también hay que analizarlos para ayudar en la toma de decisiones. Hay herramientas que nos permiten realizar estos análisis. Se pueden contar entre ellas, por ejemplo, la regresión múltiple, el análisis de componentes principales, el análisis de conglomerados y el análisis factorial. Pero la correcta aplicación de estas herramientas requiere una sólida base matemática y estadística, lo que hace difícil la comunicación entre el especialista y el gestor.

Afortunadamente, existen herramientas que nos permiten visualizar las características más importantes de los datos. Tratan de construir mapas que resuman la información de un modo gráfico. Ello hace que la comunicación entre el usuario y el analista se base principalmente en la intuición y no tanto en los conocimientos técnicos.

En este seminario se dará una introducción al análisis multivariable desde un enfoque geométrico. La herramienta estadística que se introducirá es conocida como "Escalado Multidimensional", MDS según sus siglas en inglés.

Dado un mapa no es difícil estimar la distancia entre dos puntos en él. Por ejemplo, podemos estimar que la distancia entre Logroño y Zaragoza es de unos 160 Kms mirando un mapa de España. MDS funciona al revés y utilizando un concepto más general que el de distancia: el de proximidad entre los casos. Dada una tabla de distancias o proximidades entre pares de puntos, MDS trata de construir un mapa que muestre la localización relativa de estos puntos, respetando lo más fielmente posible dicha tabla de distancias o proximidades.

Cuando la tabla de distancias cumple una serie de propiedades matemáticas, se puede usar un método analítico para resolver este problema de modo exacto. Sin embargo, en general nos conformamos con que, si la distancia o proximidad entre dos puntos es pequeña, estos puntos se sitúen cerca en el mapa; y si entre otros puntos la distancia es mayor, dichos puntos se sitúen más alejados en el mapa. Este Enfoque lo aborda el MDS ordinal, una técnica que presentaremos en este seminario.

Se hablará del concepto de proximidad y lejanía, de cómo definir proximidad y lejanía entre pares de observaciones dependiendo de la naturaleza de las escalas en que se midan sus características, de cómo construir un mapa a partir de una tabla de distancias o proximidades, de cómo decidir en cuántas dimensiones se ha de construir el mapa, de cómo evaluar la calidad de la representación, y de cómo interpretar los resultados.

Estas técnicas han sido aplicadas al análisis del fracaso empresarial, las crisis bancarias españolas, la política universitaria, la convergencia de las economías de los países de la unión europea, e incluso en análisis de la información contenida en los anuncios de oferta de empleo. Estas y otras potenciales aplicaciones se expondrán durante el seminario."