

Clase 1: Técnicas de Análisis Multivariado

Justo Andrés Manrique Urbina

24 de agosto de 2019

Usualmente, en los datos tenemos demasiadas variables: X_1, \dots, X_p . En estos datos, buscamos las interrelaciones que tienen las variables, o tal vez explicar un grupo de variables en términos de otras. Estos datos se representan en matrices cuyo marco teórico es soportado por la álgebra lineal y probabilidad. Si se estudian los datos que aparecen para cada variable X_1, \dots, X_p esto comprende el análisis multivariado. Cuando uno comienza a aprender estadístico, utiliza funciones univariadas y gráficos como histogramas o barras, para ir buscando el modelo univariado que ayude en el estudio de la variable.

El estudio multivariado inicia a partir del estudio simultáneo entre dos variables X_1, X_2 . Una de los primeros análisis es la covarianza y el coeficiente de correlación; en métodos gráficos tenemos los gráficos de dispersión (siempre y cuando sean variables numéricas).

1. Álgebra Lineal

El análisis matricial comprende una matriz tal que:

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}$$

Si $n=p$ entonces tenemos una matriz cuadrada.

La transpuesta de A es definida como A^T

La matriz identidad es definida como una matriz cuya diagonal es 1 y todos los demás elementos es 0

Una matriz es simétrica si $A^T = A$

Una matriz es idempotente si $A * A = A$

Una matriz es ortogonal si $A^T * A = I$

La traza es la suma de la suma de los elementos de la diagonal. $tr(A) = \sum_{i=1}^n a_{ii}$. La suposición es que A es cuadrada.

Determinante de una matriz es definida como $\det(A) = \sum_{\sigma \in P_n} \prod_{i=1}^n a_i \sigma_i$

La matriz inversa de A se denota de la siguiente forma: $A^{-1} \frac{C}{|A|}$

1.1. Autovectores y autovalores de una matriz

Los autovectores son, obviamente, vectores, y autovalores son escalares. Para una matriz A cuadrada, tenemos que e es un autovector de A si:

$$Ae = \lambda e$$

en donde λ es el autovalor -escalar- y e es el autovector. Generalmente se eligen autovectores unitarios. Esto quiere decir que $(\sum_{i=1}^n e_{1i}^2)^{0.5} = 1$.

Si A es simétrica los autovectores se pueden elegir de tal modo que sean unitarios, es decir $\|e_i\| = 1$ y que sean mutuamente perpendiculares. Que sean mutuamente perpendiculares significa que, para cada par de autovectores, la multiplicación de una por la otra transpuesta dan cero. Asimismo, si obtienen los autovalores de una matriz, se obtiene lo siguiente:

$$\det A = \prod_{i=1}^p \lambda_i$$

$$\text{tr}(A) = \text{tr}(\lambda) = \sum_{j=1}^p \lambda_j$$

La pregunta principal es ¿cómo encontrar los autovectores? Entonces tenemos resolver la ecuación $\det A - \lambda I = 0$.

2. Derivación

Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ y para el vector $(px1)$ la derivada es el vector columna de las derivadas parciales. Esta derivada se le conoce como la gradiente.

3. Descomposición espectral

Toda matriz simétrica se puede descomponer de la siguiente forma.

$$A = \Gamma \Lambda \Gamma^T$$

en donde Γ corresponde a la matriz formada por los autovectores de A . Esto se representaría en lo siguiente:

$$\begin{bmatrix} e_{11} & e_{21} & \cdots & e_{n1} \\ e_{12} & e_{22} & \cdots & e_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{1n} & e_{2n} & \cdots & e_{nn} \end{bmatrix} \begin{bmatrix} \Lambda_{11} & 0 & \cdots & 0 \\ 0 & \Lambda_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1p} \\ e_{21} & e_{22} & \cdots & e_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n1} & e_{n2} & \cdots & e_{np} \end{bmatrix}$$

A través de la descomposición espectral podemos hablar de la «raíz cuadrada de una matriz», **siempre y cuando la matriz siempre sea simétrica**. Formas cuadráticas... ver slide. Mardia K.V.