

Nadia Laura Kudraszow

Centro de Matemática de La Plata, Universidad Nacional de La Plata (CMaLP-UNLP) y Conicet,
Argentina

nkudraszow@mate.unlp.edu.ar

La máxima asociación entre dos variables multivariadas $\mathbf{X}$ e $\mathbf{Y}$ se define como el valor máximo que una medida de asociación bivariante entre proyecciones unidimensionales $R(\boldsymbol{\alpha}^\top \mathbf{X}, \boldsymbol{\beta}^\top \mathbf{Y})$ puede alcanzar. Tomar la correlación de Pearson como medida de asociación da como resultado el primer coeficiente de correlación canónica. Se presentará una extensión a tres o más variables multivariadas, $\mathbf{X}_1, \dots, \mathbf{X}_J$, en la que además se tiene en cuenta una red de conexiones entre las variables. Estas conexiones se codifican en una matriz de diseño $C = (c_{jk})$, donde $c_{jk} = 1$ si $\mathbf{X}_j$ y $\mathbf{X}_k$ están conectadas, y 0 en caso contrario. El objetivo es buscar la máxima asociación generalizada

$$\max_{\mathbf{w}_1, \dots, \mathbf{w}_J} \sum_{j=1}^J \sum_{k>j}^J c_{jk} g(R(\mathbf{w}_j^\top \mathbf{X}_j, \mathbf{w}_k^\top \mathbf{X}_k)),$$

donde g puede ser la función cuadrática, identidad o valor absoluto. Al igual que en el caso para dos variables, tomar la correlación de Pearson como medida de asociación da como resultado el primer coeficiente de correlación canónica generalizada de [3].

Para dos variables, [1] y [2] utilizaron diferentes medidas de asociación robusta para definir estimadores robustos de la asociación máxima. Se analizará la robustez de las medidas de máxima asociación generalizada y de los correspondientes estimadores de los coeficientes que producen la asociación máxima generalizada, extendiendo la propuesta de [1]. Se utilizarán distintas medidas robustas de asociación bivariadas como ser la de rangos de Spearman, la de Kendall y las derivadas de una matriz de dispersión bivariada. Se comparará el desempeño de la propuesta en su versión clásica y robusta a través de simulaciones de muestras gaussianas y muestras contaminadas para evaluar el comportamiento de las mismas frente a distintas desviaciones del modelo.

Trabajo en conjunto con María Victoria Fasano (Universidad Nacional de La Plata, Argentina) y Alejandra Vahnovan (Universidad Nacional de La Plata, Argentina).

Referencias

- [1] Alfons, A.; Croux, C.; Filzmoser, P.: Robust Maximum Association Estimators. In: Journal of the American Statistical Association 112 (2017), 436-445.
- [2] Branco, J. A.; Croux, C.; Filzmoser, P.; Oliveira, M. R.: Robust canonical correlations: a comparative study. In: Computational Statistics 20 (2005), 203-229.
- [3] Tenenhaus, A.; Tenenhaus, M.: Regularized Generalized Canonical Correlation Analysis. In: Psychometrika 76 (2011), Nr. 2, 257-284.