Unilateral Correspondence Analysis ー
Applied to Spanish linguistic data in space and time

Hiroto Ueda
University of Tokyo

The correspondence analysis created by Jean-Paul Benzécri is used in different scientific disciplines in general. In dialectometric studies, its validity has been demonstrated in previous works. In Japan the Quantification Method III of Chikio Hayashi is also known. The aim in common is to find the distribution of the frequencies in the two-dimensional matrix with the highest correlation coefficient, calculating the constitution of the two vectors, horizontal and vertical, which are assigned to the frequencies. As a result the frecuencies are presented in a concentrated proximity of the diagonal area, by which we can examine the relations of cases, that of attributes, and that of both cases and attributes at the same time.

The analysis of Benzécri as well as the method of Hayashi are characterized by the rearrangement of cases in rows and attributes in columns in order to obtain the best form of diagonalization in the distribution of frecuencies. However, researchers sometimes need to rearrange only the order of cases fixing the order of attributes. For example, in the historical data the chronological axis of years, decades or centuries should be maintained unchanged. We propose to derive mathematical operations to carry out the reordering of the rows of the linguistic forms provided that the columns corresponding to, for example, the chronological attributes are not to be moved.

In the presentation I will explain the mathematical derivation of the Unilateral Correspondence Analysis and its applications to Spanish dialectal data and historical documents, with demonstrations on our website. In addition I propose a method of finding the most concentrated zones within the optimally diagonalized matrix. I believe that the bilateral and unilateral correspondence analyses together with the search for the most concentrated areas will allow wide range of uses in dialectometric studies and in other fields.