

Usages linguistiques des éléments supplémentaires dans l'Analyse factorielle des correspondances

Damon Mayaffre¹, Laurent Vanni²

¹CNRS – Université Côte d'Azur – damon.mayaffre@unice.fr

²CNRS – Université Côte d'Azur – laurent.mayaffre@unice.fr

Abstract

This contribution shows the interest of supplementary variables with Correspondence Analysis (CA). From a CA vector space crossing the main morpho-syntactic categories and the French Presidents of the fifth Republic, we project the lemma “indiquer” (to indicate) as a supplementary variable. Will it be located in the “verb subspace” of the graph? And if not, what should we conclude from this counterintuitive positioning? Beyond this example, it is the linguistic homogeneity of the rows of the contingency table (words, lemmas, grammatical categories, etc.) that we question, by projecting, in more or less interpretable ways, other linguistic elements into additional elements. The contribution ends with an open discussion on the complementarity of CA and deep neural networks.

Keywords: CA, supplementary variables, illustrative elements, morpho-syntax, political discourse, Hyperbase.

Résumé

Cette contribution montre l'intérêt des éléments supplémentaires (ou éléments illustratifs) dans l'AFC. Sur un plan factoriel croisant les principales catégories morpho-syntaxiques et les présidents de la République, nous projetons en élément supplémentaire le lemme « indiquer ». Celui-ci se situera-t-il sur le graphique dans le sous-espace de la catégorie verbale ? Et si non, que faut-il conclure de ce positionnement contre-intuitif ? Au-delà de cet exemple, c'est l'homogénéité linguistique des éléments actifs du tableau de contingence, convoqués en ligne (mots, lemmes, catégories grammaticales, etc.), que nous interrogeons, en projetant, de manière plus ou moins interprétable, d'autres éléments linguistiques en éléments supplémentaires. La contribution se termine par une discussion ouverte sur la complémentarité de l'AFC et les réseaux de neurones.

Mots clés : AFC, éléments supplémentaires, éléments illustratifs, morphosyntaxe, discours politique, Hyperbase.