

ÉCONOMÉTRIE, PRÉVISION ET ANALYSE DES DONNÉES

Dominique LADIRAY
Institut National de la Statistique et des Études Économiques
15 boulevard Gabriel Péri, BP 100
92244 Malakoff Cedex
FRANCE
(dominique.ladiray@insee.fr)

1 Introduction

L'analyse des données et l'analyse des séries temporelles ont toutes deux une longue histoire mais curieusement leurs chemins se sont rarement croisés, au moins jusqu'à récemment. Dans les dix dernières années, avec la mise à disposition d'énormes ensembles de données temporelles, on a assisté à une explosion d'intérêt pour l'exploration de ces fichiers gigantesques. Des centaines de papiers ont alors présenté et diffusé des méthodes et algorithmes pour indexer, classer, discriminer, segmenter des séries temporelles. C'est à ces nouveaux liens que le présent travail s'intéresse en se concentrant sur le domaine de l'analyse de la conjoncture et la prévision à court terme des grands agrégats économiques.

Les raisons pour lesquelles l'économétrie des séries temporelles a tardé à incorporer dans ses méthodes des outils d'analyse des données sont sans doute plus subtiles et vont bien au-delà de la simple disponibilité d'importantes bases d'indicateurs économiques. Dans un article de 1989, Philip Mirowski ([31]) défend l'idée qu'il s'est écoulé une trentaine d'années avant que les concepts stochastiques de la Physique ne pénètrent en Économie. On pourrait sans doute en dire autant de l'Économétrie : le filtre de Kalman et les modèles « état-mesure » ont par exemple été développés au début des années 1960 mais n'ont été adoptés par les économètres qu'au milieu des années 80. Les retards avec les outils de l'analyse des données sont encore plus importants : il faut attendre la fin des années 80 pour voir apparaître, dans la théorie de la cointégration (Engle et Granger, 1987, [13]), les équations de l'analyse canonique et la fin des années 1990 pour voir une application de l'analyse en facteurs communs et spécifiques en analyse conjoncturelle. Bien entendu ces méthodes doivent être adaptées à la nature particulière des données temporelles et il a fallu mettre au point des algorithmes efficaces de calcul. Mais les raisons profondes du retard tiennent aux différences fondamentales de philosophies entre l'analyse des données et l'économétrie, ou plus exactement entre les analyses exploratoire et confirmatoire si chères à John Tukey ([44]). L'économètre s'accommode