
GETS MODELLING OU LASSO ? LES DIFFÉRENTES MÉTHODES DE SÉLECTION DE VARIABLES AVEC DES SÉRIES TEMPORELLES. L'EXEMPLE DES ÉTALONNAGES À L'AIDE D'ENQUÊTES DE CONJONCTURE

Clément ROUSSET, Alizé PAPP

Insee, DESE, Division Synthèse conjoncturelle

clement.rousset@insee.fr

alize.papp@ensae.fr

Mots-clés : GETS modelling, LASSO, sélection de variables, enquêtes de conjoncture, étalonnages, séries temporelles

Résumé

Les enquêtes de conjoncture fournissent des renseignements précieux pour le conjoncturiste : nombreuses, variées et rapidement disponibles, elles couvrent un large pan des économies nationales. Toutefois, face à la multiplicité des séries ainsi disponibles, allant, par exemple, de la confiance des ménages à l'anticipation de l'évolution des prix dans la construction, comment faire le tri et savoir quelles variables sont prédictives vis-à-vis d'agrégats de la comptabilité nationale et donc peuvent améliorer la prévision, et lesquelles leur sont orthogonales ?

En 2006, Éric Dubois et Emmanuel Michaux publiaient un article dans lequel ils présentaient et évaluaient deux méthodes complémentaires, à savoir la trimestrialisation des données d'enquêtes de conjoncture d'une part, et l'automatisation de la sélection des variables issues d'enquêtes de conjoncture d'autre part. Si le premier volet de leur article est très largement appliqué aujourd'hui, la sélection de variables est à la Synthèse conjoncturelle davantage artisanale. Cet article se propose de réévaluer l'opportunité d'introduire une sélection de variables automatique, et compare la méthode mise en avant par Dubois et Michaux, et la famille de méthodes qui y est liée, à savoir le *General-to-Specific modelling* (GETS) et les méthodes de Lasso (*Least Absolute Shrinkage and Selection Operator*).

Pour ce faire, différentes simulations ont été effectuées, en faisant varier l'intensité du bruit, sa distribution et la proportion de variables informatives. L'usage du LASSO et du *GETS modelling* permet ensuite de comparer les avantages et inconvénients de chaque méthode, et de trancher sur l'opportunité de leur utilisation pour la division de la Synthèse conjoncturelle. En particulier, on a quatre souhaits pour cette méthode de sélection – souhaits qui fournissent chacun un critère d'évaluation :

- *l'écoute précise*, que la méthode distingue le signal du bruit, et ne rejette pas les bonnes variables,
- *la discrimination*, le rejet des variables qui ne sont pas informatives,
- *la prédiction*, que le modèle retenu soit prédictif,
- *la stabilité*, une méthode qui non seulement satisfasse les trois critères précédents, mais aussi le fasse fréquemment (à l'opposé d'une méthode par exemple excellente pour la moitié des résultats mais très mauvaise pour les autres) – on pourra ainsi préférer une méthode un peu moins bonne mais plus stable.

On commencera par présenter ces deux approches relativement différentes de la sélection de variables, à savoir d'une part le *GETS modelling*, issu des économètres de la LSE, et d'autre part le LASSO, en soulignant leurs hypothèses et paramétrages respectifs. Après cela, on présentera les résultats des simulations, au regard des trois critères sus-mentionnés. Enfin, l'on s'attachera à un cas précis d'usage à la Synthèse conjoncturelle, et aux éventuelles améliorations de prévision qui en ont découlé. Par-delà l'usage des enquêtes de conjoncture, la question fondamentale qui se pose est alors celle de l'équilibre possible entre automatisation aveugle et « agnostique » et méthode artisanale mais interprétative.

Bibliographie

- [1] Éric Dubois, Emmanuel Michaux, « Étalonnages à l'aide d'enquêtes de conjoncture : de nouveaux résultats », *Economie & prévision* 2006/1 (no 172), p. 11-28.
- [2] Sucarrat, Genaro, « Automated financial multi-path GETS modelling », 2009
- [3] R. Tibshirani, « Regression Shrinkage and Selection via the Lasso », in : *Journal of the Royal Statistical Society. Series B (Methodological)* 58.1 (1996), p. 267–288.
- [4] Julia Campos, Neil R. Ericsson, David F. Hendry, « General-to-specific Modelling: An Overview and Selected Bibliography », 2005, Board of Governors of the Federal Reserve System, International Finance Discussion Papers
- [5] Ivan Savin, « A comparative study of the Lasso-type and heuristic model selection methods », 2010