

8 - Séries Longues

Désaisonnalisation avec JDemetra+

Anna Smyk & Tanguy Barthélémy (Insee)



Sommaire I

① Problèmes d'estimation

② Raccourcissement de la période d'estimation

Section 1

Problèmes d'estimation

Séries Longues

Quand et pourquoi les problèmes d'estimation apparaissent ?

Deux cas :

- rupture (plusieurs régimes) : mélanges de signaux (périodiques)
- série "juste" trop longue (approx. plus de 20 ans) : estimations sous-optimales en fin de période (le plus regardé)

Premier cas : sous périodes évidentes

Dans la suite nous nous concentrons sur le 2ème cas: le processus générateur de données évolue en même temps que les institutions et les comportements (ex: travail du dimanche, consommation de produits frais...)

Rigidités de l'estimation ?

Si on utilise X13-Arima

Pré-ajustement fait "à coefficients fixes" sur l'intégralité de la période : rigide

- effets de calendrier
- détection de points aberrants (outliers)
- structure d'auto-corrélation (modèles Arima)

Décomposition : plus souple

- X-13 : moyennes mobiles (fenêtres mobiles, souple)
- X-13 : utilisation d'une prévision en fin de série avec modèle Arima (rigide)

Cas favorables / défavorables

D'un point de vue de l'estimation, en dehors de l'évolution intrinsèque des données

Favorable (peu impactant de ne pas couper)

- pas d'effet cjo

Défavorable

- effets cjo notables

Couper est recommandé par les Guidelines dans tous les cas.

Section 2

Raccourcissement de la période d'estimation

Mesurer l'influence de la période d'estimation (1/2)

Comparer l'estimation sur période optimale (O) (8-12 ans) et l'estimation sur période longue (L), avec des indicateurs du type:

- moyenne des écarts en niveau :

$$EN(Y_t) = \frac{1}{T} \sum_{k=1}^T \left| \frac{cvs_k^L - cvs_k^O}{cvs_k^L} \right| \times 100$$

Mesurer l'influence de la période d'estimation (2/2)

- moyenne des écarts en taux d'évolution

$$ETC(Y_t) = \frac{1}{T-1} \sum_{k=2}^T \left| \frac{cvs_k^L - cvs_{k-1}^L}{cvs_{k-1}^L} - \frac{cvs_k^O - cvs_{k-1}^O}{cvs_{k-1}^O} \right| \times 100$$

- proportion de points avec sens d'évolution contraires

$$CohS(Y_t) = \frac{1}{T-1} \sum_{k=2}^T I_{(cvs_k^L - cvs_{k-1}^L)(cvs_k^O - cvs_{k-1}^O) < 0} \times 100$$

(Voir exemple dans article JMS 2018, cf. infra)

Restreindre ces indicateurs aux 3 ou 5 dernières années

Recommandations

- couper en sous périodes qui se chevauchent (longueur d'estimation 8-12 ans)
- couper en dehors des crises
- raccorder et ne reviser que la dernière sous période (différents raccordements possibles, en général on accole, cf article JMS 2018)
- publier les cvs-cjo révisées sur une période plus courte que la dernière période d'estimation
- revoir les sous-périodes tous les 5-8 ans

Pour plus de détails voir: Guidelines 2024, articles Pham, Quartier-La-Tente JMS 2018 et Smyk JMS 2022 (in /Series_Longues dans répertoire GitHub)

Exemple

Indice de la production industrielle

- séries débutant en général en 1990
- en 2019 coupure: début estimation 2004, raccord en 2011, révisions en publication à partir de 2012
- en 2024 coupure: debut estimation 2012, raccord 2017*, révisions en publication à partir de 2018 (?)

(*en theorie)

Mais aussi Douanes, Enquêtes de conjoncture (comportent des séries commençant en 1979 (!), désaisonnalisées avec Tramo-Seats, mais pas d'effet cjo...)