

# Extraction de facteurs communs par filtre de Kalman : Application à des données macroéconomiques françaises

Projet d'Économétrie Financière

Morgan Jowitt, Pierre Liberge, Aya Mokthar, Gaétan Dumas

Mai 2025

## **Abstract**

Ce projet a pour objectif de reproduire et d'illustrer les résultats méthodologiques présentés dans l'article "Factor extraction using Kalman Filter and smoothing" de Poncela, Ruiz et Miranda (2021), en les appliquant à un nouveau jeu de données macroéconomiques françaises. L'article propose une approche robuste d'extraction de facteurs communs latents à partir de séries temporelles multivariées, via un Modèle à Facteurs Dynamiques (DFM) couplé à un filtre et lisseur de Kalman. À travers une implémentation en Python, nous mettons en œuvre cette approche sur un panel macroéconomique français entre 1975 et 2025, et comparons les résultats à une Analyse en Composantes Principales (ACP). Nous montrons que l'ACP, bien que statique, offre de meilleures corrélations avec les variables observées que le modèle dynamique de Kalman dans ce cas d'étude.

# 1 Introduction

Dans un contexte économique où les décideurs doivent suivre un grand nombre d'indicateurs macroéconomiques, les modèles à facteurs dynamiques (DFM) permettent de résumer efficacement l'information en extrayant quelques facteurs communs non observés. L'article "Factor Extraction using Kalman Filter and Smoothing" (Poncela et al., 2021) propose une méthodologie complète pour extraire ces facteurs en combinant une représentation en espace d'état, le filtre de Kalman et l'algorithme EM pour estimer les paramètres du modèle.

L'objectif de ce projet est de reproduire la méthodologie de l'article sur un nouveau jeu de données économiques françaises. Pour cela, nous suivons les mêmes étapes : préparation des données, estimation des paramètres par maximum de vraisemblance via l'algorithme EM, extraction des facteurs à l'aide du filtre et lisseur de Kalman, puis comparaison avec l'approche par composantes principales (PCA).

## 2 Modèle DFM et Filtrage de Kalman

Le modèle dynamique à facteur unique est représenté par :

$$Y_t = \Lambda F_t + \varepsilon_t, \quad F_t = \Phi F_{t-1} + \eta_t$$

où :

- $Y_t$  est un vecteur de variables observées,
- $F_t$  est un facteur latent,
- $\Lambda$  est la matrice de chargement,
- $\varepsilon_t \sim \mathcal{N}(0, H)$  et  $\eta_t \sim \mathcal{N}(0, Q)$  sont les bruits.

## 3 Méthodologie

Les données couvrent la période 1975-2025 et comprennent : le climat des affaires, le taux de chômage mensualisé, l'indice des prix à la consommation, la production industrielle et le taux d'intérêt à 10 ans.

Les données ont été nettoyées par interpolation et outlier-clipping via l'IQR, puis normalisées. Trois approches ont été implémentées :

- DFM via `statsmodels.DynamicFactor`,
- DFM personnalisé avec `MLEModel`,
- ACP via `sklearn.decomposition.PCA`.

## 4 Résultats Empiriques

### 4.1 Performance des Modèles

### 4.2 Visualisation des Facteurs

Lecture de la figure :

	Kalman (DFM)	Kalman (Custom)	ACP
Log-vraisemblance	-1867.64	-	-
AIC	3773.28	-	-
Variance expliquée (facteur 1)	-	-	52.1%
Corrélation moyenne (facteur 1)	0.548	-	0.676

Table 1: Comparaison des performances des méthodes sur 604 observations (1975–2025)

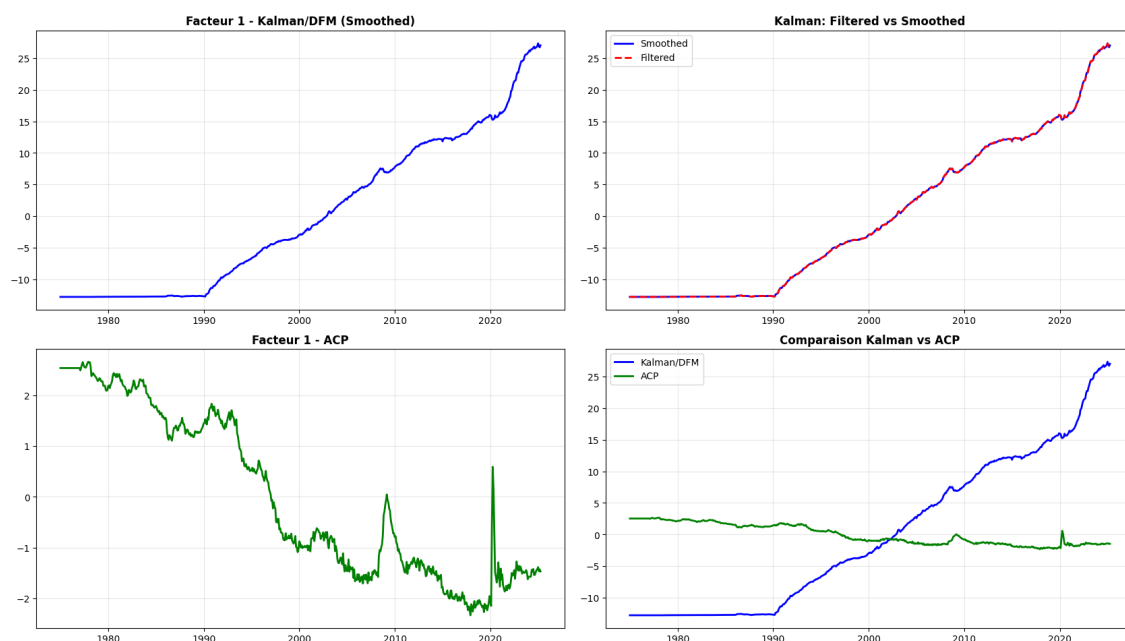


Figure 1: Comparaison des facteurs extraits (Kalman Smoothed & Filtered vs ACP)

- **Haut gauche :** Le facteur 1 extrait via Kalman (smoothed) montre une forte dynamique monotone croissante sur plus de trois décennies, traduisant une tendance structurelle longue.
- **Haut droite :** Le facteur lissé (bleu) et le facteur filtré (rouge pointillé) coïncident fortement, ce qui suggère un bon ajustement du modèle.
- **Bas gauche :** Le facteur 1 de l'ACP est centré, plus volatil, avec des pics nets autour de 2009 et 2020, cohérents avec les crises financières et COVID.
- **Bas droite :** La comparaison directe montre un décalage structurel : Kalman suit une trajectoire de niveau tandis que l'ACP reste stationnaire autour de zéro, mais reflète mieux les chocs de court terme.

### 4.3 Corrélations avec les Variables Observées

#### Kalman (Facteur 1) :

- Indice des prix à la consommation : +1.000
- Taux à 10 ans : -0.915

- Production industrielle : +0.513
- Taux de chômage : +0.263
- Climat des affaires : +0.048

#### ACP (Composante 1) :

- Taux à 10 ans : +0.952
- Indice des prix à la consommation : +0.875
- Production industrielle : +0.772
- Taux de chômage : +0.515
- Climat des affaires : +0.267

**Corrélation entre les deux approches :** Le facteur Kalman (1) et la composante ACP (1) sont fortement corrélés entre eux : 0.879, ce qui indique qu'ils capturent une dynamique latente similaire, bien que de forme différente.

**Interprétation :** Le facteur issu du Kalman semble capter des tendances de long terme dominées par l'inflation et les taux (composante monétaire/structurelle), tandis que l'ACP capte plus de volatilité à court terme et reflète mieux l'activité réelle.

**Recommandation finale :** Au vu des corrélations plus homogènes et de la variance expliquée, l'ACP est ici plus efficace statistiquement, bien que moins structurelle. Le filtre de Kalman reste utile pour explorer les dynamiques de fond et construire des prévisions économiques.

## 5 Interprétation à la Lumière de la Littérature

Le facteur extrait par Kalman est très fortement corrélé à l'inflation et inversement au taux d'intérêt à 10 ans, ce qui indique qu'il capture probablement une dynamique monétaire ou nominale. En revanche, le climat des affaires est peu corrélé, ce qui fragilise une interprétation purement réelle.

L'ACP produit un facteur plus équilibré, expliquant mieux la variance conjointe des séries économiques. Cela corrobore l'idée que l'ACP, bien que statique, reste robuste lorsque la structure des séries est bien alignée linéairement.

## 6 Conclusion

Dans notre étude, le filtre de Kalman n'a pas surpassé l'ACP en termes de performance explicative, malgré ses avantages théoriques. Toutefois, il offre une perspective temporelle précieuse et peut être amélioré via des extensions (facteurs multiples, modèles bayésiens, changements de régimes).

## References