



Institut national de la statistique
et des études économiques

Mesurer pour comprendre

Estimation en temps réel de la tendance-cycle : Apport de l'utilisation des filtres asymétriques dans la détection des points de retournement

ALAIN QUARTIER-LA-TENTE
Division Études macroéconomiques
20 septembre 2022

Sommaire

1. Introduction

2. Méthodes étudiées

3. Comparaison des méthodes

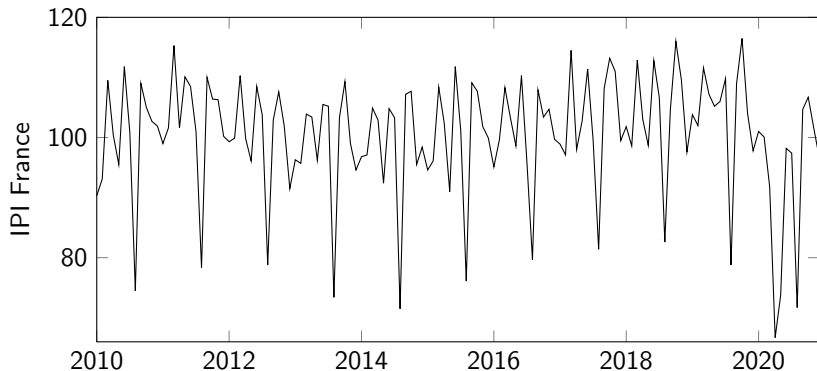
4. Extensions

5. Conclusion

La tendance-cycle (1)

X_t (ex : IPI France) se décompose en plusieurs composantes inobservées :

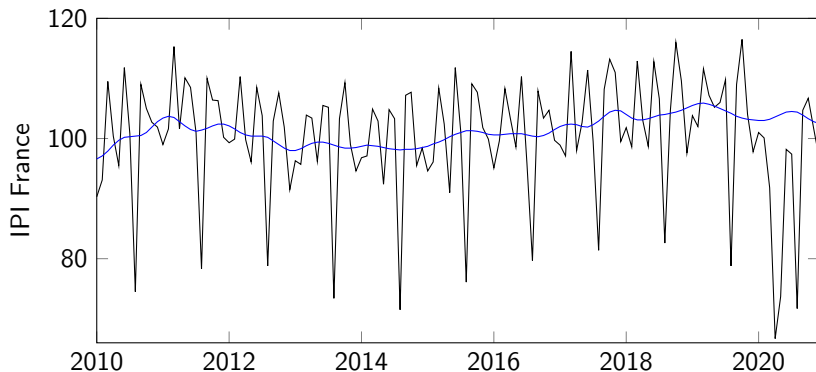
$$X_t = \quad \quad \quad (\text{décomposition additive})$$



La tendance-cycle (1)

X_t (ex : IPI France) se décompose en plusieurs composantes inobservées :

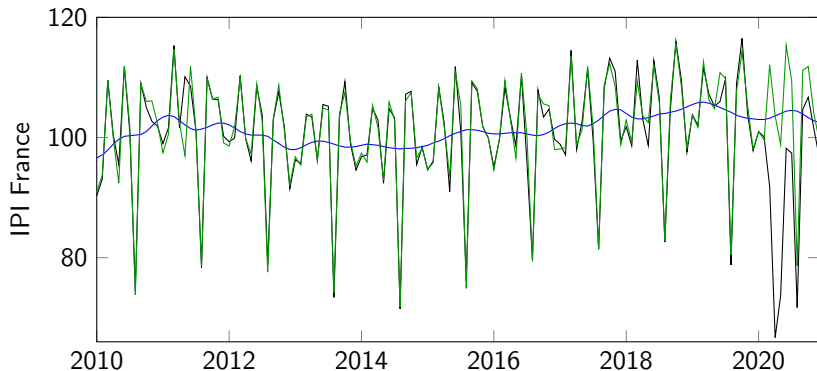
$$X_t = \underbrace{TC_t}_{\text{tendance-cycle}} \quad (\text{décomposition additive})$$



La tendance-cycle (1)

X_t (ex : IPI France) se décompose en plusieurs composantes inobservées :

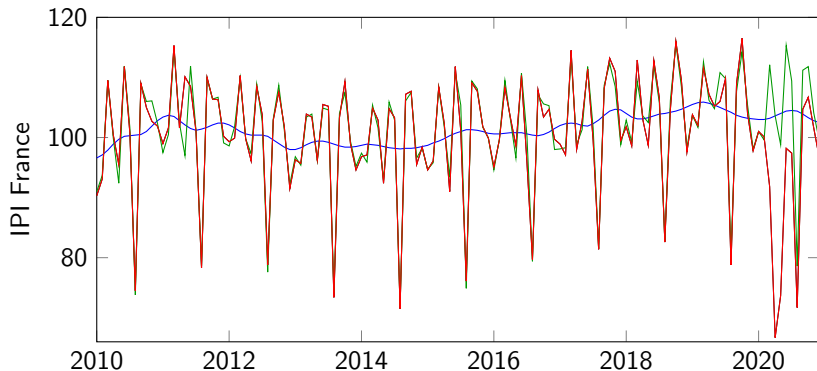
$$X_t = \underbrace{TC_t}_{\text{tendance-cycle}} + \underbrace{S_t}_{\text{saisonnalité}} \quad (\text{décomposition additive})$$



La tendance-cycle (1)

X_t (ex : IPI France) se décompose en plusieurs composantes inobservées :

$$X_t = \underbrace{TC_t}_{\text{tendance-cycle}} + \underbrace{S_t}_{\text{saisonnalité}} + \underbrace{I_t}_{\text{irrégulier}} \quad (\text{décomposition additive})$$

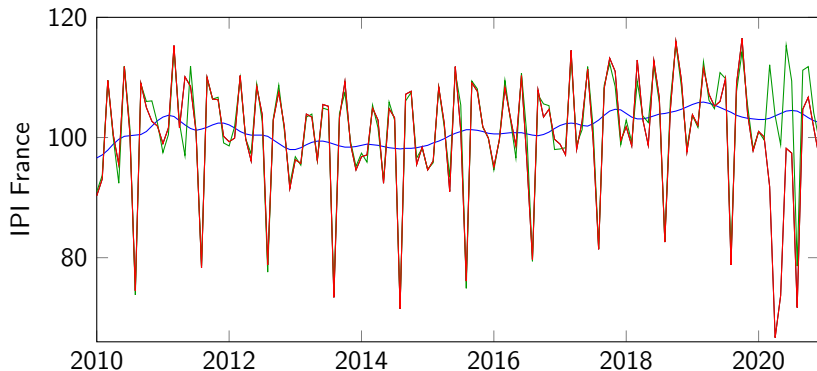


La tendance-cycle (1)

X_t (ex : IPI France) se décompose en plusieurs composantes inobservées :

$$X_t = \underbrace{TC_t}_{\text{tendance-cycle}} + \underbrace{S_t}_{\text{saisonnalité}} + \underbrace{I_t}_{\text{irrégulier}} \quad (\text{décomposition additive})$$

tendance et cycle ici estimés *simultanément*



La tendance-cycle (2)

Pour l'analyse conjoncturelle, on étudie généralement des séries désaisonnalisées

$$X_t - S_t = TC_t + I_t$$

La tendance-cycle (2)

Pour l'analyse conjoncturelle, on étudie généralement des séries désaisonnalisées

$$X_t - S_t = TC_t + I_t$$

La présence de l'irrégulier peut rendre l'interprétation difficile, un lissage supplémentaire peut être utilisé :

$$(X_t - S_t) - I_t = TC_t$$

La tendance-cycle (2)

Pour l'analyse conjoncturelle, on étudie généralement des séries désaisonnalisées

$$X_t - S_t = TC_t + I_t$$

La présence de l'irrégulier peut rendre l'interprétation difficile, un lissage supplémentaire peut être utilisé :

$$(X_t - S_t) - I_t = TC_t$$

Tendance-cycle publiée par peu d'instituts (Insee, ONS, Statistics Canada, ABS) mais volonté de faire des bonnes pratiques au niveau européen (Destatis).

La tendance-cycle (2)

Pour l'analyse conjoncturelle, on étudie généralement des séries désaisonnalisées

$$X_t - S_t = TC_t + I_t$$

La présence de l'irrégulier peut rendre l'interprétation difficile, un lissage supplémentaire peut être utilisé :

$$(X_t - S_t) - I_t = TC_t$$

Tendance-cycle publiée par peu d'instituts (Insee, ONS, Statistics Canada, ABS) mais volonté de faire des bonnes pratiques au niveau européen (Destatis).

Critères importants de la tendance-cycle :

1. minimiser les révisions
2. minimiser le nombre de faux points de retournement
3. détecter correctement et *rapidement* les (bons) points de retournement

Estimations de la TC et moyennes mobiles (1)



objectif différent de celui des méthodes d'analyse des cycles économiques
 TC_t généralement estimée sur une série *sans* saisonnalité

Estimations de la TC et moyennes mobiles (1)



objectif différent de celui des méthodes d'analyse des cycles économiques

TC_t généralement estimée sur une série *sans* saisonnalité

Méthode de décomposition X-13ARIMA une des plus utilisées : études de méthodes non-paramétriques pour estimer TC_t

Estimations de la TC et moyennes mobiles (1)



objectif différent de celui des méthodes d'analyse des cycles économiques

TC_t généralement estimée sur une série *sans* saisonnalité

Méthode de décomposition X-13ARIMA une des plus utilisées : études de méthodes non-paramétriques pour estimer TC_t

Moyennes mobiles (ou *filtres linéaires*) omniprésents dans l'extraction de la tendance-cycle et la désaisonnalisation (e.g. : X-13ARIMA) :

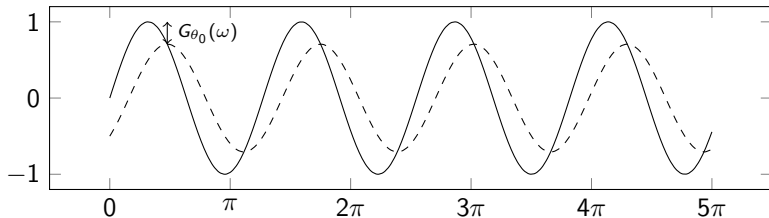
$$M_{\theta}(X_t) = \sum_{k=-p}^{+f} \theta_k X_{t+k}$$

Estimations de la TC et moyennes mobiles (1)

Appliquer M_θ sur $X_t = e^{-i\omega t}$ va avoir deux effets :

$$M_\theta X_t = \sum_{k=-p}^{+f} \theta_k e^{-i\omega(t+k)} = \left(\sum_{k=-p}^{+f} \theta_k e^{-i\omega k} \right) \cdot X_t = G_\theta(\omega) e^{-i\Phi_\theta(\omega)} X_t$$

1. Multiplier le niveau par $G_\theta(\omega)$ (*gain*)

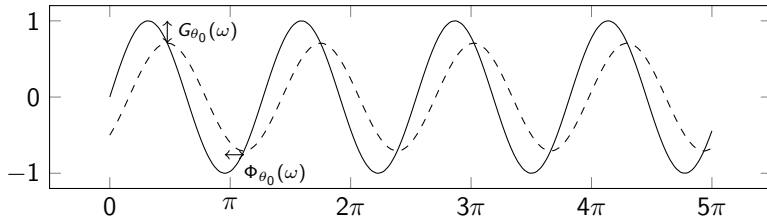


Estimations de la TC et moyennes mobiles (1)

Appliquer M_θ sur $X_t = e^{-i\omega t}$ va avoir deux effets :

$$M_\theta X_t = \sum_{k=-p}^{+f} \theta_k e^{-i\omega(t+k)} = \left(\sum_{k=-p}^{+f} \theta_k e^{-i\omega k} \right) \cdot X_t = G_\theta(\omega) e^{-i\Phi_\theta(\omega)} X_t$$

1. Multiplier le niveau par $G_\theta(\omega)$ (*gain*)
2. Créer un *déphasage* $\Phi_\theta(\omega)/\omega$: affecte détection des points de retournement



Estimations de la TC et moyennes mobiles (2)

➡ Généralement, utilisation de filtres *symétriques* ($p = f$ et $\theta_{-i} = \theta_i$)

$$M_{\theta}(X_t) = \sum_{k=-p}^{+p} \theta_k X_{t+k}, \quad \text{avec } \theta_{-i} = \theta_i$$

Estimations de la TC et moyennes mobiles (2)

➡ Généralement, utilisation de filtres *symétriques* ($p = f$ et $\theta_{-i} = \theta_i$)

$$M_{\theta}(X_t) = \sum_{k=-p}^{+p} \theta_k X_{t+k}, \quad \text{avec } \theta_{-i} = \theta_i$$

➡ Pour l'estimation en **temps réel**, utilisation de filtres *asymétriques* ($f < p$) \implies révision et détection avec retard des points de retournement (*déphasage*)

$$\text{ex : } M_{\theta}(X_t) = \sum_{k=-p}^0 \theta_k X_{t+k}$$

Estimations de la TC et moyennes mobiles (2)

➡ Généralement, utilisation de filtres *symétriques* ($p = f$ et $\theta_{-i} = \theta_i$)

$$M_{\theta}(X_t) = \sum_{k=-p}^{+p} \theta_k X_{t+k}, \quad \text{avec } \theta_{-i} = \theta_i$$

➡ Pour l'estimation en *temps réel*, utilisation de filtres *asymétriques* ($f < p$) \implies révision et détection avec retard des points de retournement (*déphasage*)

$$\text{ex : } M_{\theta}(X_t) = \sum_{k=-p}^0 \theta_k X_{t+k}$$

Solution classique : prolonger la série par prévision et utiliser filtre symétrique

➡ revient à utiliser des filtres asymétriques optimisés avec certains critères

➡ sous-optimal pour séries très variables

Objectifs

Objectifs cette étude :

- Étudier et comparer des approches récentes pour l'extraction de la tendance-cycle en temps réel : Régression polynomiale locale (Proietti et Luati 2008); RKHS (Dagum et Bianconcini 2016); Optimisation sous contrainte d'une somme pondérée de critères (Grun-Rehomme *et ali* 2018, Wildi et McElroy, 2019)


Objectifs

Objectifs cette étude :

- Étudier et comparer des approches récentes pour l'extraction de la tendance-cycle en temps réel : Régression polynomiale locale (Proietti et Luati 2008); RKHS (Dagum et Bianconcini 2016); Optimisation sous contrainte d'une somme pondérée de critères (Grun-Rehomme *et ali* 2018, Wildi et McElroy, 2019)
- Montrer qu'il est possible d'établir une théorie générale englobant toutes ces méthodes

Objectifs

Objectifs cette étude :

- Étudier et comparer des approches récentes pour l'extraction de la tendance-cycle en temps réel : Régression polynomiale locale (Proietti et Luati 2008); RKHS (Dagum et Bianconcini 2016); Optimisation sous contrainte d'une somme pondérée de critères (Grun-Rehomme *et ali* 2018, Wildi et McElroy, 2019)
- Montrer qu'il est possible d'établir une théorie générale englobant toutes ces méthodes
- Présenter le package  `rjdfilters`
<https://github.com/palatej/rjdfilters>

Sommaire

1. Introduction

2. Méthodes étudiées

2.1 Filtre symétrique

2.2 Polynômes Locaux

2.3 Filtres et Reproducing Kernel Hilbert Space (RKHS)

2.4 Minimisation sous contrainte : FST et ATS

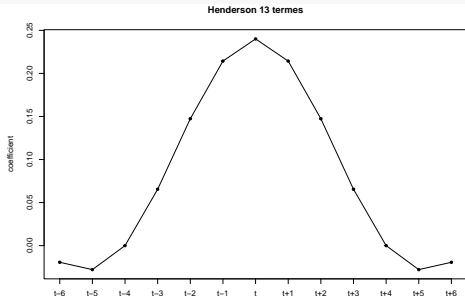
3. Comparaison des méthodes

4. Extensions

5. Conclusion

Moyenne mobile symétrique d'Henderson

```
library(rjdfilters)
f <- lp_filter(6, kernel = "Henderson")
plot_coef(f, q = 6, legend = FALSE,
          main="Henderson 13 termes")
```



MM Henderson (utilisé dans X-13ARIMA) largement répandue pour estimer TC_t

MM Henderson préserve les tendances polynomiales de degré 3 et minimise le critère de "lissage" ($\sum (\nabla^3 \theta_i)^2$)

Sur séries mensuelles : MM de 13 termes généralement

Polynômes Locaux : `rjdfilters::lp_filter()`

Hypothèse : $y_t = \mu_t + \varepsilon_t$ avec $\varepsilon_t \stackrel{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$

μ_t localement approchée par un polynôme de degré d :

$$\forall j \in \llbracket -h, h \rrbracket : y_{t+j} = m_{t+j} + \varepsilon_{t+j}, \quad m_{t+j} = \sum_{i=0}^d \beta_i j^i$$

Polynômes Locaux : `rjdfilters::lp_filter()`

Hypothèse : $y_t = \mu_t + \varepsilon_t$ avec $\varepsilon_t \stackrel{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$

μ_t localement approchée par un polynôme de degré d :

$$\forall j \in \llbracket -h, h \rrbracket : y_{t+j} = m_{t+j} + \varepsilon_{t+j}, \quad m_{t+j} = \sum_{i=0}^d \beta_i j^i$$

Estimation en utilisant les WLS avec *noyaux* : $\hat{\beta} = (X' K X)^{-1} X' K y$ et

$$\hat{m}_t = \hat{\beta}_0 = w' y = \sum_{j=-h}^h w_j y_{t-j} \Rightarrow \text{équivalent à une moyenne mobile symétrique}$$

➔ Filtre de Henderson avec $d = 3$ et noyau spécifique.

Filtres asymétriques : `rjdfilters::lp_filter()`

1. Même méthode mais moins de données (DAF) \iff minimiser les révisions sous mêmes contraintes polynomiales
 - ➔ **sans biais** mais **beaucoup de variance**
 - ➔ utilisé dans STL

Filtres asymétriques : `rjdfilters::lp_filter()`

1. Même méthode mais moins de données (DAF) \iff minimiser les révisions sous mêmes contraintes polynomiales
 - ➔ **sans biais** mais **beaucoup de variance**
 - ➔ utilisé dans STL
 2. Minimisation des révisions sous contraintes polynomiales :
 - 2.1 *Linear-Constant* (LC) : y_t linéaire and v reproduit les constantes (*Musgrave*)
 - 2.2 *Quadratic-Linear* (QL) : y_t quadratique et v reproduit droites
 - 2.3 *Cubic-Quadratic* (CQ) : y_t cubique et v reproduit tendances quadratiques
- ➔ Filtres asymétriques v dépendent de "IC-Ratio"

Filtres asymétriques : `rjdfilters::lp_filter()`

1. Même méthode mais moins de données (DAF) \iff minimiser les révisions sous mêmes contraintes polynomiales

➡ **sans biais** mais **beaucoup de variance**

➡ utilisé dans STL

2. Minimisation des révisions sous contraintes polynomiales :

2.1 *Linear-Constant* (LC) : y_t linéaire and v reproduit les constantes (*Musgrave*)

2.2 *Quadratic-Linear* (QL) : y_t quadratique et v reproduit droites

2.3 *Cubic-Quadratic* (CQ) : y_t cubique et v reproduit tendances quadratiques

➡ Filtres asymétriques v dépendent de “IC-Ratio”



modèles simples facilement interprétables



Déphasage non contrôlé ➡ méthode étendue dans `rjdfilters::lp_filter()`

Filtres asymétriques : `rjdfilters::lp_filter()`

1. Même méthode mais moins de données (DAF) \iff minimiser les révisions sous mêmes contraintes polynomiales

➡ **sans biais** mais **beaucoup de variance**

➡ utilisé dans STL

2. Minimisation des révisions sous contraintes polynomiales :

2.1 *Linear-Constant* (LC) : y_t linéaire and v reproduit les constantes (*Musgrave*)

2.2 *Quadratic-Linear* (QL) : y_t quadratique et v reproduit droites

2.3 *Cubic-Quadratic* (CQ) : y_t cubique et v reproduit tendances quadratiques

➡ Filtres asymétriques v dépendent de “IC-Ratio”



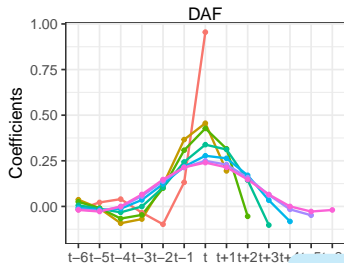
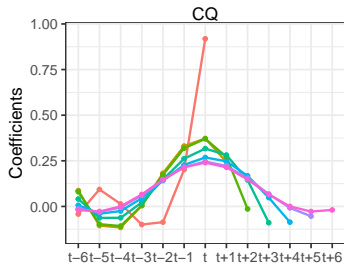
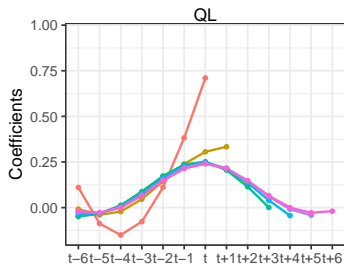
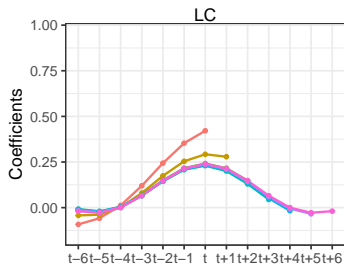
modèles simples facilement interprétables



Déphasage non contrôlé ➡ méthode étendue dans `rjdfilters::lp_filter()`

🖥 Visualisation <https://aqlt.shinyapps.io/FiltersProperties/>

Coefficients



Filtres RKHS : `rjdfilters::rkhs_filter()`

- Utilisation de la théorie des RKHS pour approcher le filtre d'Henderson
- Avec K_p une **fonction de noyau** définie sur $[-1, 1]$, le filtre symétrique :

$$\forall j \in \llbracket -h, h \rrbracket : w_j = \frac{K_p(j/b)}{\sum_{i=-h}^h K_p(i/b)}$$

➔ avec $b = h + 1$ et K_p spécifique on retrouve le filtre d'Henderson

- Pour les filtres asymétriques :

$$\forall j \in \llbracket -h, q \rrbracket : w_{a,j} = \frac{K_p(j/b)}{\sum_{i=-h}^q K_p(i/b)}$$

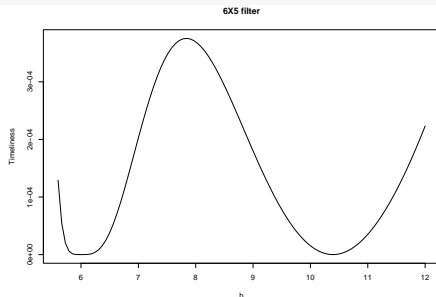
➔ b choisit par optimisation, e.g. minimisant les révisions $(b_{q,r})$, les révisions liées à la fonction de gain $(b_{q,G})$ et celles liées au déphasage $(b_{q,\varphi})$

Filtres asymétriques



Plusieurs extremum

```
library(rjdfilters)
fun <- rkhs_optimization_fun(horizon = 6,
                             leads = 5, degree = 3,
                             asymmetricCriterion = "Timeliness")
plot(fun, 5.6, 12, xlab = "b",
      ylab = "Timeliness", main = "6X5 filter")
```



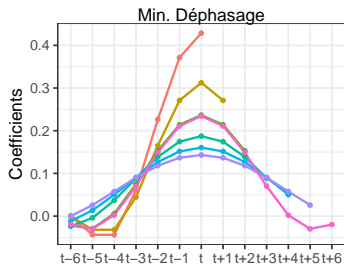
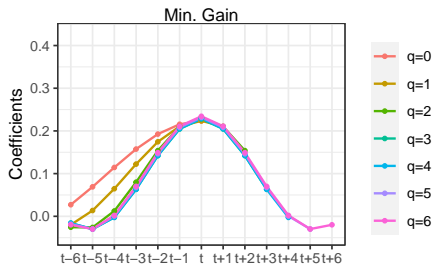
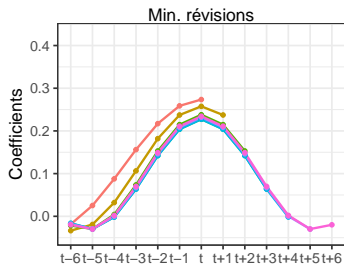
```
rkhs_optimal_bw()
```

##	q=0	q=1	q=2	q=3	q=4	q=5
##	6.0000	6.0000	6.3875	8.1500	9.3500	6.0000



Méthode
généralisable à des
filtres avec fréquences
irrégulières

Coefficients



Approche FST : `rjdfilters::fst_filter()`

Minimisation sous contrainte d'une somme pondérée de 3 critères :

$$\begin{cases} \min_{\theta} & J(\theta) = \alpha F_g(\theta) + \beta S_g(\theta) + \gamma T_g(\theta) \\ s.c. & C\theta = a \end{cases}$$

F_g fidélité (*fidelity*, réduction de variance $\sum_{k=-p}^{+f} \theta_k^2$), S_g lissage (*smoothness*, critère d'Henderson $\sum_j (\nabla^3 \theta_j)^2$), T_g temporalité (*timeliness*, déphasage $\int_0^{2\pi/12} \rho_\theta(\omega) \sin(\varphi_\theta(\omega))^2 d\omega$)

Approche FST : `rjdfilters::fst_filter()`

Minimisation sous contrainte d'une somme pondérée de 3 critères :

$$\begin{cases} \min_{\theta} & J(\theta) = \alpha F_g(\theta) + \beta S_g(\theta) + \gamma T_g(\theta) \\ \text{s.c.} & C\theta = a \end{cases}$$

F_g fidélité (*fidelity*, réduction de variance $\sum_{k=-p}^{+f} \theta_k^2$), S_g lissage (*smoothness*, critère d'Henderson $\sum_j (\nabla^3 \theta_j)^2$), T_g temporalité (*timeliness*, déphasage $\int_0^{2\pi/12} \rho_{\theta}(\omega) \sin(\varphi_{\theta}(\omega))^2 d\omega$)

- 😊 Solution unique
- 😊 Filtres asymétriques indépendants des données et du filtre symétrique
- 😞 Poids non normalisés

Approche ATS `rjdfilters::dfa_filter()` (1)

Décomposition de l'EQM :

$$\begin{aligned}\mathbb{E} [(y_t - \hat{y}_t)^2] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\Gamma_s(\omega) - \Gamma_\theta(\omega)|^2 h(\omega) d\omega \\ &= \frac{1}{2\pi} \times 2 \times \int_0^{\pi} |\Gamma_s(\omega) - \Gamma_\theta(\omega)|^2 h(\omega) d\omega\end{aligned}$$

et

$$\begin{aligned}|\Gamma_s(\omega) - \Gamma_\theta(\omega)|^2 &= \rho_s(\omega)^2 + \rho_\theta(\omega)^2 + \\ &\quad 2\rho_s(\omega)\rho_\theta(\omega)(1 - \cos(\varphi_s(\omega) - \varphi_\theta(\omega))) \\ &= (\rho_s(\omega) - \rho_\theta(\omega))^2 + \\ &\quad 4\rho_s(\omega)\rho_\theta(\omega)\sin^2\left(\frac{\varphi_s(\omega) - \varphi_\theta(\omega)}{2}\right)\end{aligned}$$

Approche ATS `rjdfilters::dfa_filter()` (2)

Ce qui conduit à

$$A_w = 2 \int_0^{\omega_1} (\rho_s(\omega) - \rho_\theta(\omega))^2 h(\omega) d\omega$$

$$T_w = 8 \int_0^{\omega_1} \rho_s(\lambda) \rho_\theta(\lambda) \sin^2 \left(\frac{\varphi_\theta(\omega)}{2} \right) h(\omega) d\omega$$

$$S_w = 2 \int_{\omega_1}^{\pi} (\rho_s(\omega)^2 - \rho_\theta(\omega))^2 h(\omega) d\omega$$




$$R_w = 8 \int_{\omega_1}^{\pi} \rho_s(\lambda) \rho_\theta(\lambda) \sin^2 \left(\frac{\varphi_\theta(\omega)}{2} \right) h(\omega) d\omega$$

Minimisation d'une somme pondérée de 3 critères :

$$\mathcal{M}(\vartheta_1, \vartheta_2) = \vartheta_1 T_w(\theta) + \vartheta_2 S_w(\theta) + (1 - \vartheta_1 - \vartheta_2) A_w(\theta)$$

\implies minimisation sous contraintes linéaires avec $h(\omega) = 1$

Approche ATS `rjdfilters::dfa_filter()` (3)

-  Poids ont un sens
-  Résidus pas toujours négligeables
-  Pas unicité de la solution

Difficulté du choix des poids

Comment choisir les poids dans FST et AST (DFA) ?

- Minimiser que la *timeliness*? introduit trop de variance

Difficulté du choix des poids

Comment choisir les poids dans FST et AST (DFA) ?

- Minimiser que la *timeliness*? introduit trop de variance
- Minimiser les révisions? on néglige le déphasage

Difficulté du choix des poids

Comment choisir les poids dans FST et AST (DFA) ?

- Minimiser que la *timeliness*? introduit trop de variance
- Minimiser les révisions? on néglige le déphasage
- Garder les poids qui donnent des MM de meilleure qualité que les autres selon tous les critères? Rien de satisfaisant

Difficulté du choix des poids

Comment choisir les poids dans FST et AST (DFA) ?

- Minimiser que la *timeliness*? introduit trop de variance
- Minimiser les révisions? on néglige le déphasage
- Garder les poids qui donnent des MM de meilleure qualité que les autres selon tous les critères? Rien de satisfaisant
- Faire quadrillage du plan et une analyse empirique du déphasage?
Difficile à analyser -> utiliser des méthodes de réduction de dimension ?

Sommaire

1. Introduction

2. Méthodes étudiées

3. Comparaison des méthodes

3.1 Méthodologie

3.2 Application sur séries simulées

3.3 Un exemple : série des ventes au détail des États-Unis (en log)

4. Extensions

5. Conclusion

Méthodologie

Comparaison des différentes méthodes sur séries simulées (avec 3 niveaux de variabilité) et séries réelles :

1. Estimation de la tendance-cycle à chaque date en utilisant les différentes méthodes et un filtre symétrique de 13 termes

Méthodologie

Comparaison des différentes méthodes sur séries simulées (avec 3 niveaux de variabilité) et séries réelles :

1. Estimation de la tendance-cycle à chaque date en utilisant les différentes méthodes et un filtre symétrique de 13 termes
2. À chaque date, estimation des points de retournement :
 - redressements : $y_{t-3} \geq y_{t-2} \geq y_{t-1} < y_t \leq y_{t+1}$
 - ralentissements : $y_{t-3} \leq y_{t-2} \leq y_{t-1} > y_t \geq y_{t+1}$

Déphasage = temps nécessaire pour détecter le bon point de retournement
sans révision

Méthodologie

Comparaison des différentes méthodes sur séries simulées (avec 3 niveaux de variabilité) et séries réelles :

1. Estimation de la tendance-cycle à chaque date en utilisant les différentes méthodes et un filtre symétrique de 13 termes
2. À chaque date, estimation des points de retournement :
 - redressements : $y_{t-3} \geq y_{t-2} \geq y_{t-1} < y_t \leq y_{t+1}$
 - ralentissements : $y_{t-3} \leq y_{t-2} \leq y_{t-1} > y_t \geq y_{t+1}$

Déphasage = temps nécessaire pour détecter le bon point de retournement
sans révision

3. Calcul des révisions avec deux critères :

$$\mathbb{E} \left[\left| \frac{y_{t|t+q} - y_{t|last}}{y_{t|last}} \right| \right] \quad \text{et} \quad \mathbb{E} \left[\left| \frac{y_{t|t+q} - y_{t|t+q+1}}{y_{t|t+q+1}} \right| \right]$$

Séries simulées (1)

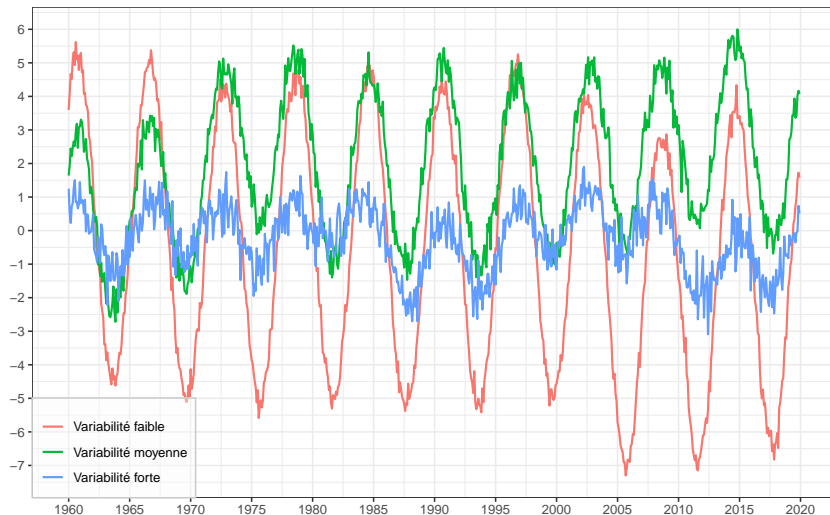
De façon similaire à Darne et Dagum (2009), on simule $y_t = C_t + T_t + I_t$ entre janvier 1960 et décembre 2020 :

- $C_t = \rho[\cos(2\pi t/\lambda) + \sin(2\pi t/\lambda)]$, $\lambda = 72$ (cycles de 6 ans, 19 points de retournement détectables)
- $T_t = T_{t-1} + \nu_t$ avec $\nu_t \sim \mathcal{N}(0, \sigma_\nu^2)$, $\sigma_\nu = 0,08$
- $I_t = e_t$ avec $e_t \sim \mathcal{N}(0, \sigma_e^2)$

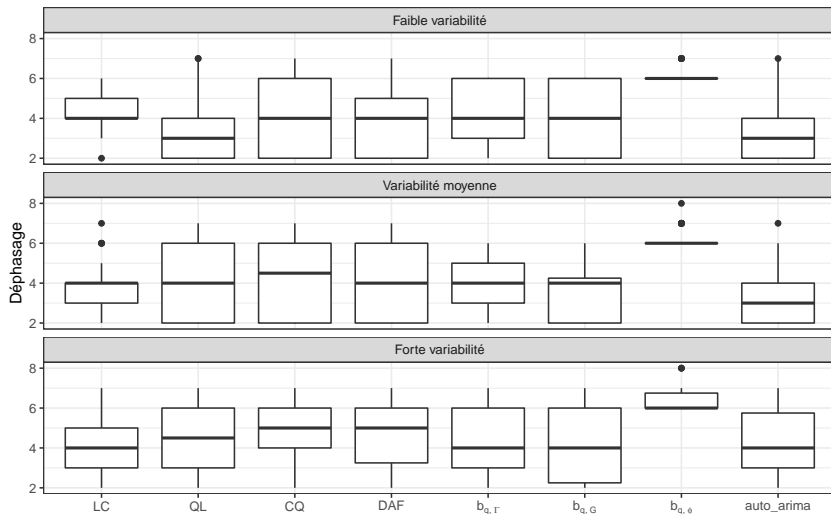
Niveau de variabilité :

- variabilité faible (rapport signal/bruit fort) : $\sigma_e^2 = 0,2$ et $\rho = 3,0, 3,5$ ou $4,0$ ($0,9 \geq \text{I-C ratio} \geq 0,7$)
- variabilité moyenne (rapport signal/bruit moyen) : $\sigma_e^2 = 0,3$ et $\rho = 1,5, 2,0$ ou $3,0$ ($2,3 \geq \text{I-C ratio} \geq 1,4$)
- variabilité forte (rapport signal/bruit élevé) : $\sigma_e^2 = 0,4$ et $\rho = 0,5, 0,7$ ou $1,0$ ($8,9 \geq \text{I-C ratio} \geq 5,2$)

Séries simulées (2)



Résultats sur le déphasage (séries simulées)



Médiane des révisions (séries simulées)

Pour les séries à variabilité **moyenne** :

Méthode	$q = 0$	$q = 1$	$q = 2$	$q = 3$	$q = 4$	$q = 5$
MAE entre q^e et la dernière estimation						
LC	0,21	0,10	0,03	0,03	0,03	0,01
QL (rel)	1,6	1,0	1,3	1,5	1,3	1,1
CQ (rel)	2,2	1,3	4,2	3,3	2,1	1,6
DAF (rel)	2,3	1,5	4,9	3,5	2,2	1,5
$b_{q,\Gamma}$ (rel)	3,1	2,3	1,1	3,6	3,5	3,9
$b_{q,G}$ (rel)	4,1	4,0	1,1	3,6	3,5	4,0
$b_{q,\varphi}$ (rel)	1,5	1,1	1,0	1,8	2,7	8,7
ARIMA (rel)	1,0	1,0	1,1	1,1	1,1	1,0
MAE entre q^e et la $q + 1^e$ estimation						
LC	0,19	0,10	0,02	0,01	0,07	0,01
QL (rel)	1,6	43,2	0,1	3,1	0,9	1,1
CQ (rel)	2,3	0,2	4,3	7,3	1,4	1,6
DAF (rel)	3,5	2,6	4,6	12,9	1,3	1,5
$b_{q,\Gamma}$ (rel)	2,1	2,9	3,7	0,3	16,2	3,9
$b_{q,G}$ (rel)	3,7	4,7	4,3	0,5	17,0	4,0
$b_{q,\varphi}$ (rel)	1,2	1,4	3,5	5,3	0,7	8,7
ARIMA (rel)	1,1	1,3	0,8	1,8	2,3	1,0

Médiane des révisions (séries simulées)

Pour les séries à variabilité **moyenne** :

Méthode	$q = 0$	$q = 1$	$q = 2$	$q = 3$	$q = 4$	$q = 5$
MAE entre q^e et la dernière estimation						
LC	0,21	0,10	0,03	0,03	0,03	0,01
QL (rel)	1,6	1,0	1,3	1,5	1,3	1,1
CQ (rel)	2,2	1,3	4,2	3,3	2,1	1,6
DAF (rel)	2,3	1,5	4,9	3,5	2,2	1,5
$b_{q,\Gamma}$ (rel)	3,1	2,3	1,1	3,6	3,5	3,9
$b_{q,G}$ (rel)	4,1	4,0	1,1	3,6	3,5	4,0
$b_{q,\varphi}$ (rel)	1,5	1,1	1,0	1,8	2,7	8,7
ARIMA (rel)	1,0	1,0	1,1	1,1	1,1	1,0
MAE entre q^e et la $q + 1^e$ estimation						
LC	0,19	0,10	0,02	0,01	0,07	0,01
QL (rel)	1,6	43,2	0,1	3,1	0,9	1,1
CQ (rel)	2,3	0,2	4,3	7,3	1,4	1,6
DAF (rel)	3,5	2,6	4,6	12,9	1,3	1,5
$b_{q,\Gamma}$ (rel)	2,1	2,9	3,7	0,3	16,2	3,9
$b_{q,G}$ (rel)	3,7	4,7	4,3	0,5	17,0	4,0
$b_{q,\varphi}$ (rel)	1,2	1,4	3,5	5,3	0,7	8,7
ARIMA (rel)	1,1	1,3	0,8	1,8	2,3	1,0

Médiane des révisions (séries simulées)

Pour les séries à variabilité **moyenne** :

Méthode	$q = 0$	$q = 1$	$q = 2$	$q = 3$	$q = 4$	$q = 5$
MAE entre q^e et la dernière estimation						
LC	0,21	0,10	0,03	0,03	0,03	0,01
QL (rel)	1,6	1,0	1,3	1,5	1,3	1,1
CQ (rel)	2,2	1,3	4,2	3,3	2,1	1,6
DAF (rel)	2,3	1,5	4,9	3,5	2,2	1,5
$b_{q,\Gamma}$ (rel)	3,1	2,3	1,1	3,6	3,5	3,9
$b_{q,G}$ (rel)	4,1	4,0	1,1	3,6	3,5	4,0
$b_{q,\varphi}$ (rel)	1,5	1,1	1,0	1,8	2,7	8,7
ARIMA (rel)	1,0	1,0	1,1	1,1	1,1	1,0
MAE entre q^e et la $q + 1^e$ estimation						
LC	0,19	0,10	0,02	0,01	0,07	0,01
QL (rel)	1,6	43,2	0,1	3,1	0,9	1,1
CQ (rel)	2,3	0,2	4,3	7,3	1,4	1,6
DAF (rel)	3,5	2,6	4,6	12,9	1,3	1,5
$b_{q,\Gamma}$ (rel)	2,1	2,9	3,7	0,3	16,2	3,9
$b_{q,G}$ (rel)	3,7	4,7	4,3	0,5	17,0	4,0
$b_{q,\varphi}$ (rel)	1,2	1,4	3,5	5,3	0,7	8,7
ARIMA (rel)	1,1	1,3	0,8	1,8	2,3	1,0

Médiane des révisions (séries simulées)

Pour les séries à variabilité **moyenne** :

Méthode	$q = 0$	$q = 1$	$q = 2$	$q = 3$	$q = 4$	$q = 5$
MAE entre q^e et la dernière estimation						
LC	0,21	0,10	0,03	0,03	0,03	0,01
QL (rel)	1,6	1,0	1,3	1,5	1,3	1,1
CQ (rel)	2,2	1,3	4,2	3,3	2,1	1,6
DAF (rel)	2,3	1,5	4,9	3,5	2,2	1,5
$b_{q,\Gamma}$ (rel)	3,1	2,3	1,1	3,6	3,5	3,9
$b_{q,G}$ (rel)	4,1	4,0	1,1	3,6	3,5	4,0
$b_{q,\varphi}$ (rel)	1,5	1,1	1,0	1,8	2,7	8,7
ARIMA (rel)	1,0	1,0	1,1	1,1	1,1	1,0
MAE entre q^e et la $q + 1^e$ estimation						
LC	0,19	0,10	0,02	0,01	0,07	0,01
QL (rel)	1,6	43,2	0,1	3,1	0,9	1,1
CQ (rel)	2,3	0,2	4,3	7,3	1,4	1,6
DAF (rel)	3,5	2,6	4,6	12,9	1,3	1,5
$b_{q,\Gamma}$ (rel)	2,1	2,9	3,7	0,3	16,2	3,9
$b_{q,G}$ (rel)	3,7	4,7	4,3	0,5	17,0	4,0
$b_{q,\varphi}$ (rel)	1,2	1,4	3,5	5,3	0,7	8,7
ARIMA (rel)	1,1	1,3	0,8	1,8	2,3	1,0

Médiane des révisions (séries simulées)

Pour les séries à variabilité **faible** :

Méthode	$q = 0$	$q = 1$	$q = 2$	$q = 3$	$q = 4$	$q = 5$
MAE entre qième et la dernière estimation						
LC	0,1	0,1	0,0	0,0	0,0	0,0
QL (rel)	1,7	1,1	1,2	1,1	1,0	0,9
CQ (rel)	2,1	1,7	4,9	3,1	1,8	1,1
DAF (rel)	2,2	2,0	5,8	3,6	2,0	1,3
$b_{q,\Gamma}$ (rel)	6,6	5,0	1,3	6,1	7,3	7,5
$b_{q,G}$ (rel)	8,8	8,9	1,1	6,2	7,3	7,5
$b_{q,\varphi}$ (rel)	2,6	1,9	1,3	1,7	3,3	9,8
ARIMA (rel)	1,4	1,2	1,6	1,6	1,3	1,0
MAE entre qième et la $q + 1$ième estimation						
LC	0,1	0,0	0,0	0,1	0,0	0,0
QL (rel)	1,4	1,3	0,0	1,8	1,5	0,9
CQ (rel)	2,3	0,3	5,4	5,2	2,0	1,1
DAF (rel)	1,6	0,7	5,8	7,8	2,2	1,3
$b_{q,\Gamma}$ (rel)	3,4	22,8	8,7	0,5	12,8	7,5
$b_{q,G}$ (rel)	13,7	10,2	11,6	0,8	9,1	7,5
$b_{q,\varphi}$ (rel)	1,5	3,0	2,3	2,8	1,7	9,8
ARIMA (rel)	2,5	1,2	0,3	2,1	1,7	1,0

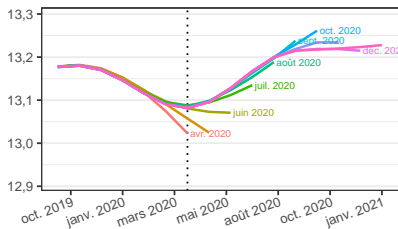
Médiane des révisions (séries simulées)

Pour les séries à variabilité **forte** :

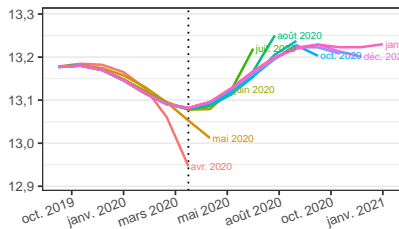
Méthode	$q = 0$	$q = 1$	$q = 2$	$q = 3$	$q = 4$	$q = 5$
MAE entre qieme et la dernière estimation						
LC	1,6	0,6	0,3	0,2	0,3	0,1
QL (rel)	1,8	1,2	1,1	1,5	1,4	1,4
CQ (rel)	2,1	2,2	5,0	4,1	2,7	2,1
DAF (rel)	2,4	2,3	4,6	4,3	2,5	2,3
$b_{q,\Gamma}$ (rel)	1,0	1,1	1,2	1,0	1,1	1,4
$b_{q,G}$ (rel)	1,2	1,2	1,1	1,0	1,1	1,4
$b_{q,\varphi}$ (rel)	1,2	1,2	1,2	1,9	3,4	10,7
ARIMA (rel)	1,0	1,2	1,0	1,1	0,8	0,9
MAE entre qieme et la $q + 1$ieme estimation						
LC	0,7	0,4	0,1	0,0	0,2	0,1
QL (rel)	2,0	1,5	0,1	3,5	2,5	1,4
CQ (rel)	1,7	0,2	19,1	9,2	1,4	2,1
DAF (rel)	11,0	1,4	4,2	9,3	1,6	2,3
$b_{q,\Gamma}$ (rel)	1,0	1,3	1,3	0,2	0,9	1,4
$b_{q,G}$ (rel)	0,9	1,5	8,5	0,2	0,8	1,4
$b_{q,\varphi}$ (rel)	1,2	1,6	3,3	4,3	1,0	10,7
ARIMA (rel)	32,8	1,4	0,8	1,6	0,9	0,9

Estimations successives de la tendance-cycle (1)

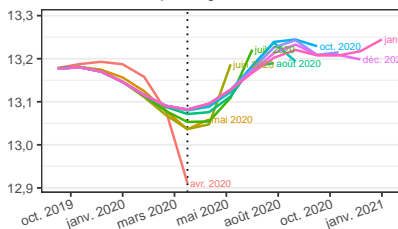
Linear – Constant (LC)
Déphasage de 3 mois



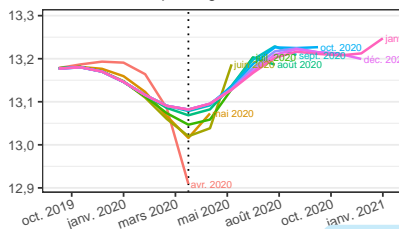
Quadratic – Linear (QL)
Déphasage de 2 mois



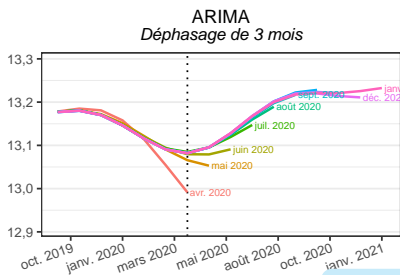
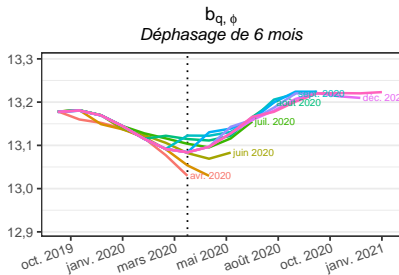
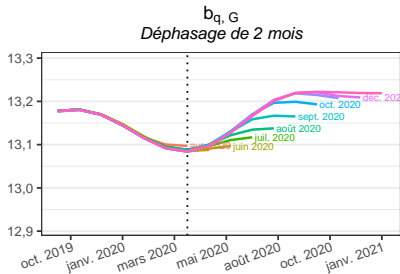
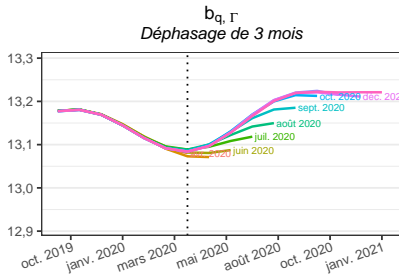
Cubic – quadratic (CQ)
Déphasage de 2 mois



DAF
Déphasage de 2 mois



Estimations successives de la tendance-cycle (2)



Prévisions implicites

Fonction `rjdfilters::implicit_forecast`

$$\forall q, \underbrace{\sum_{i=-h}^0 v_i y_i + \sum_{i=1}^h v_i y_i^*}_{\text{lissage par } v \text{ de la série prolongée}} = \underbrace{\sum_{i=-h}^0 w_i^q y_i + \sum_{i=1}^h w_i^q y_i^*}_{\text{lissage par } w^q \text{ de la série prolongée}} \quad \text{avec } \forall i > q, w_i^q = 0$$

Ce qui est équivalent à :

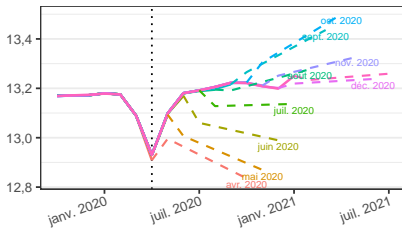
$$\forall q, \sum_{i=1}^h (v_i - w_i^q) y_i^* = \sum_{i=-h}^0 (w_i^q - v_i) y_i.$$

Matriciellement :

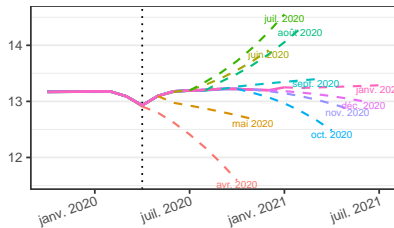
$$\begin{pmatrix} v_1 & v_2 & \cdots & v_h \\ v_1 - w_1^1 & v_2 & \cdots & v_h \\ \vdots & \vdots & \cdots & \vdots \\ v_1 - w_1^{h-1} & v_2 - w_2^{h-1} & \cdots & v_h \end{pmatrix} \begin{pmatrix} y_1^* \\ \vdots \\ y_h^* \end{pmatrix} = \begin{pmatrix} w_{-h}^0 - v_{-h} & w_{-(h-1)}^0 - v_{-(h-1)} & \cdots & w_0^0 - v_0 \\ w_{-h}^1 - v_{-h} & w_{-(h-1)}^1 - v_{-(h-1)} & \cdots & w_0^1 - v_0 \\ \vdots & \vdots & \cdots & \vdots \\ w_{-h}^{h-1} - v_{-h} & w_{-(h-1)}^{h-1} - v_{-(h-1)} & \cdots & w_0^{h-1} - v_0 \end{pmatrix} \begin{pmatrix} y_{-h} \\ \vdots \\ y_0 \end{pmatrix}.$$

Prévisions implicites (1)

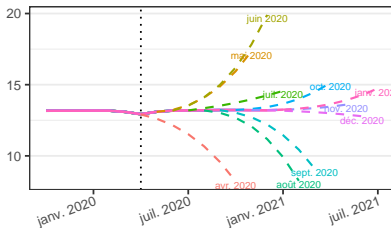
Linear – Constant (LC)
Déphasage de 3 mois



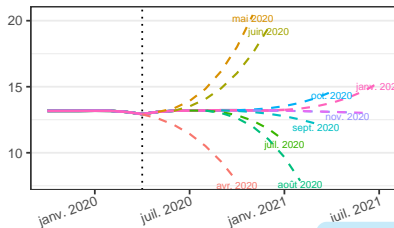
Quadratic – Linear (QL)
Déphasage de 2 mois



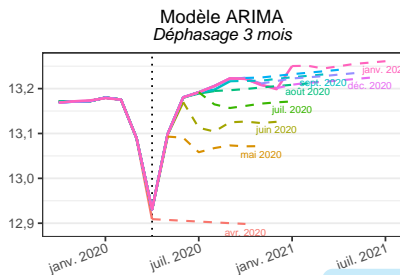
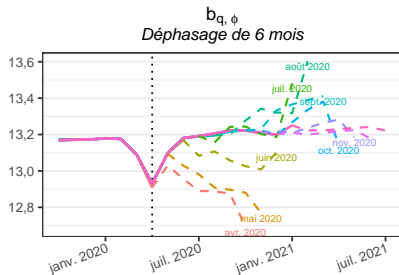
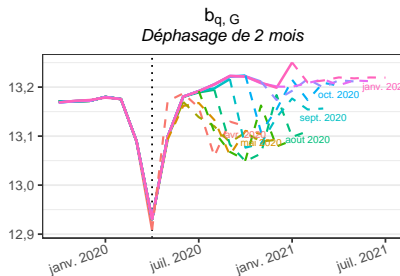
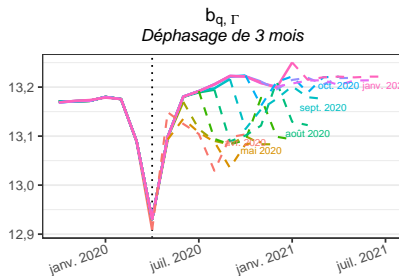
Cubic – quadratic (CQ)
Déphasage de 2 mois



DAF
Déphasage de 2 mois



Prévisions implicites (2)



Sommaire

1. Introduction

2. Méthodes étudiées

3. Comparaison des méthodes

4. Extensions

4.1 Choix de la fenêtre

4.2 Paramétrisation locale des méthodes polynomiales

5. Conclusion

Combien de termes utiliser les MM asymétriques ? (1)

Actuellement on utilise toujours autant de points dans le passé (6) que la MM symétriques pour les estimations intermédiaires : hypothèse raisonnable ? Faudrait-il utiliser plus ou moins de points dans le passé ?

Critères classiques : validation croisée, CP-Mallow, AIC, Rice-T :

$$CV(\hat{\mu}) = \frac{1}{n-2h} \sum_{t=h+1}^{n-h} \frac{(y_t - \hat{\mu}_t)^2}{(1-w_0)^2}$$

$$CP(\hat{\mu}) = \frac{1}{\sigma^2} \sum_{t=h+1}^{n-h} (y_t - \hat{\mu}_t)^2 - (n-2h)(1-2w_0)$$

Mais en général leur minimisation ne donne pas de bon résultats (critères peu discriminants)

Pistes à explorer :

Combien de termes utiliser les MM asymétriques ? (2)

1. Méthodes plus complexes de sélection de la fenêtre (e.g. Fan et Gijbels 1992)
2. Méthode des plus proches voisins : utiliser toujours le même nombre de points (e.g. toujours 13 points)

Estimation de la pente (1)

Régression non paramétrique : $y_i = \mu(x_i) + \varepsilon_i$ avec ε_i un terme d'erreur.

Avec Taylor, pour tout point x_0 , si μ est différentiable d fois, alors :

$$\forall x : \mu(x) = \mu(x_0) + \mu'(x_0)(x - x_0) + \cdots + \frac{\mu^{(d)}(x_0)}{d!}(x - x_0)^d + R_d(x),$$

Régression polynomiale

Hypothèse : $y_t = \mu_t + \varepsilon_t$ avec $\varepsilon_t \stackrel{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$

μ_t localement approchée par un polynôme de degré d :

$$\forall j \in \llbracket -h, h \rrbracket : y_{t+j} = \sum_{i=0}^d \beta_i j^i + \varepsilon_{t+j}$$

Estimation de la pente (2)

Estimation en utilisant les WLS avec *noyaux* : $\hat{\beta} = (X'KX)^{-1}X'Ky$ et

$$\hat{m}_t = \hat{\beta}_0 = e_1' \hat{\beta} = w' y = \sum_{j=-h}^h w_j y_{t-j} \text{ avec } e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Et de la même façon :

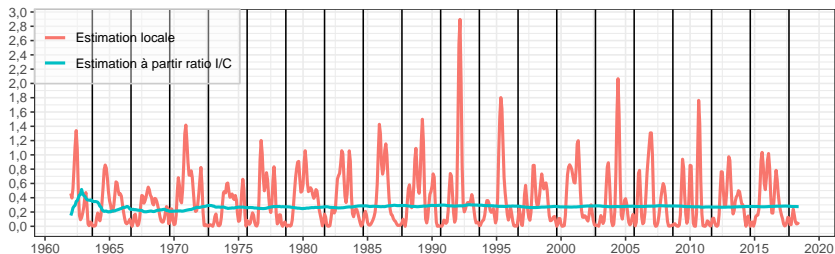
$$\begin{cases} \hat{\beta}_1 = \widehat{\mu'(t)} = e_2' \hat{\beta} & (\neq \widehat{\mu(t)})' \\ \hat{\beta}_2 = \widehat{\mu''(t)} = e_3' \hat{\beta} \end{cases}$$

Dans la méthode LC, en fin de période on suppose :

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t \text{ avec } \varepsilon_t \sim \mathcal{N}(0, \sigma^2)$$

Estimation de la pente (3)

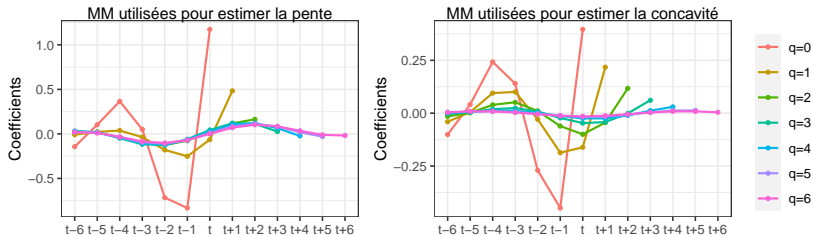
Filtres asymétriques dépendent du ratio $|\beta_1/\sigma|$ qui est toujours supposé constant : peu de sens au niveau global, notamment dans les périodes de points de retournement ($\beta_1 \simeq 0$)



Idée : paramétrisation locale

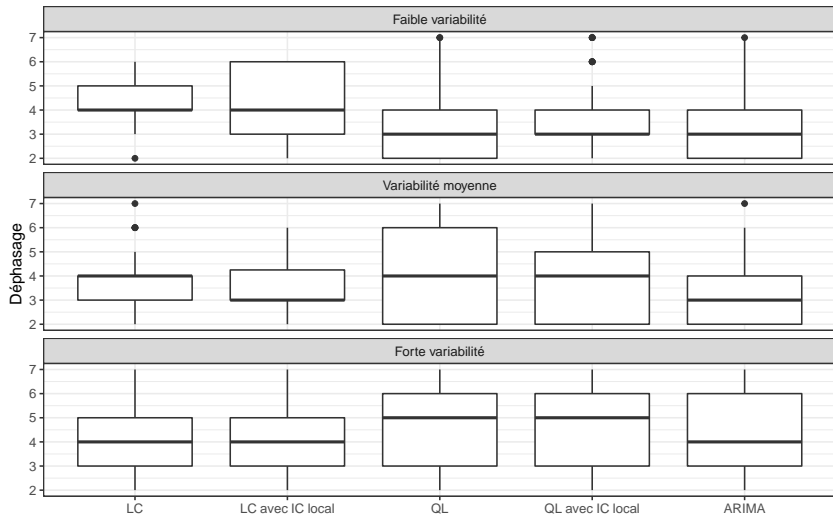
$$\begin{cases} \hat{\sigma}^2 = \frac{1}{n-2h} \sum_{t=h+1}^{n-h} \frac{(y_t - \hat{\mu}_t)^2}{1 - 2w_0^2 + \sum w_i^2} \\ \beta_1 \text{ et } \beta_2 \text{ estimés par MM (DAF par simplification)} \end{cases}$$

Estimation de la pente (4)



Rmq : il y a (encore) de fortes révisions entre la première et deuxième estimation, on pourrait utiliser méthode QL pour avoir les estimateurs de la pente

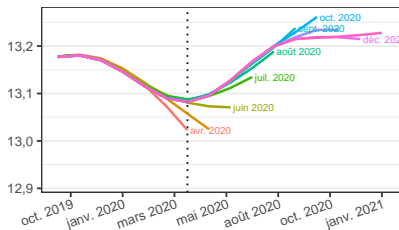
Résultats (1)



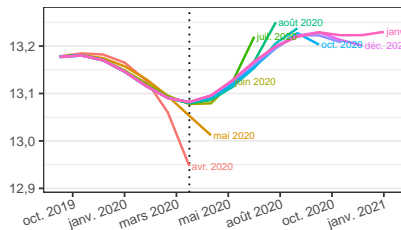
Résultats (2)

Sur la série étudiée, les résultats sont très proches :

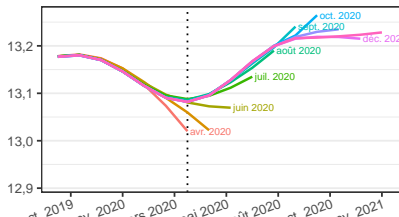
Linear – Constant (LC)
Déphasage de 3 mois



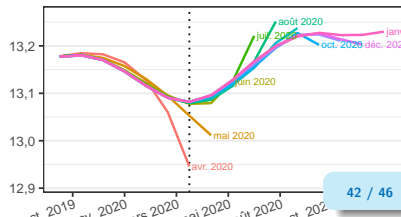
Quadratic – Linear (QL)
Déphasage de 2 mois



LC avec IC local
Déphasage de 3 mois

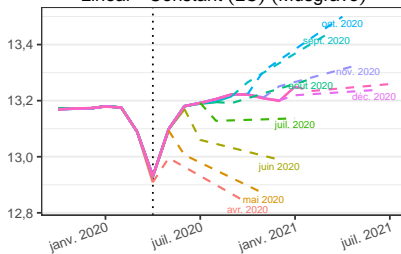


QL avec IC local
Déphasage de 2 mois

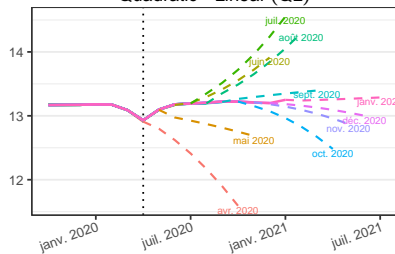


Résultats (3)

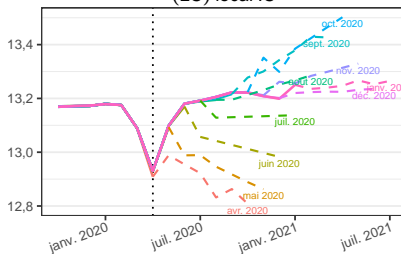
Linear – Constant (LC) (Musgrave)



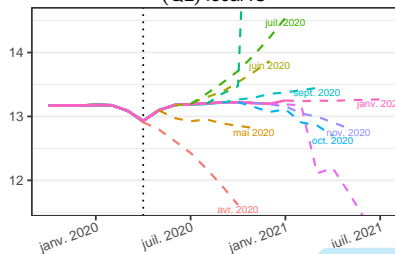
Quadratic – Linear (QL)



(LC) local IC



(QL) local IC



Sommaire

1. Introduction

2. Méthodes étudiées

3. Comparaison des méthodes

4. Extensions

5. Conclusion

5.1 Conclusion

Conclusion

- Dans la construction des filtres asymétriques :
1. on peut se restreindre à ceux qui conservent les polynômes de degré au plus 1 (et exclure les filtres QL, CQ et DAF)

Conclusion

- Dans la construction des filtres asymétriques :
 1. on peut se restreindre à ceux qui conservent les polynômes de degré au plus 1 (et exclure les filtres QL, CQ et DAF)
 2. on peut utiliser le filtre LC pour les estimations proches de l'estimation finale

Conclusion

- Dans la construction des filtres asymétriques :
 1. on peut se restreindre à ceux qui conservent les polynômes de degré au plus 1 (et exclure les filtres QL, CQ et DAF)
 2. on peut utiliser le filtre LC pour les estimations proches de l'estimation finale
- Dans certains cas des méthodes alternatives à la prévision ARIMA peuvent être utilisées ➡ `rjdfilters` peut aider à comparer les résultats (`rjdfilters::x11()` pour les intégrer dans X-11)


What next?

- Etudes sur d'autres méthodes comme Vasyechko et Grun-Rehorme (2014) ou Feng et Schäfer (2021) ou l'extension des méthodes polynomiales avec T_g

What next?

- Etudes sur d'autres méthodes comme Vasyechko et Grun-Rehorme (2014) ou Feng et Schäfer (2021) ou l'extension des méthodes polynomiales avec T_g
- Impact des points atypiques ? quid des méthodes robustes ?

Merci pour votre attention

Package  :

 palatej/rjdfilters

Exemples d'applications : https://aqlt.github.io/slides/2022%20-%2006%20-%20SAPW/Slides_AQLT.pdf