Proietti, Luati, Real Time Estimation in Local Polynomial Regression, with Application to Trend-Cycle Analysis (2008)

Alain Quartier-la-Tente
16/06/2020
Ensae — 2019-2020

Introduction

- Problème récurrent en macroéconomie : étude des points de retournement et évaluations de tendances sous-jacentes ➔ méthodes d'extraction du signal
- Problème commun à ces méthodes : fiabilité de l'estimation en fin de période
- Objectif : étude des propriétés des estimations en temps réel par des polynômes locaux

Cas général

On suppose que l'on peut décomposer notre série $y_t = \mu_t + \varepsilon_t$ que μ_t le signal/tendance et $\varepsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$.

μ_t est localement approché par un polynôme de degré d :

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

Cas général

On suppose que l'on peut décomposer notre série $y_t = \mu_t + \varepsilon_t$ que μ_t le signal/tendance et $\varepsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$.

μ_t est localement approché par un polynôme de degré d :

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

Pour estimer les coefficients il faut $2h \geq d$, on utilise les moindres carrés pondérés (WLS) en minimisant :

$$S(\hat\beta_0, \dots, \hat\beta_d) = \sum_{j=-h}^{h} \kappa_j (y_{t+j} - \hat\beta_0 - \hat\beta_1 j - \dots - \hat\beta_d j^d)^2$$

Cas général

On suppose que l'on peut décomposer notre série $y_t = \mu_t + \varepsilon_t$ que μ_t le signal/tendance et $\varepsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$.

μ_t est localement approché par un polynôme de degré d :

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

Pour estimer les coefficients il faut $2h \geq d$, on utilise les moindres carrés pondérés (WLS) en minimisant :

$$S(\hat{\beta}_0, \dots, \hat{\beta}_d) = \sum_{j=-h}^{h} \kappa_j (y_{t+j} - \hat{\beta}_0 - \hat{\beta}_1 j - \dots - \hat{\beta}_d j^d)^2$$

L'estimateur est $\hat{\beta} = (X'KX)^1 X'Ky$ et :

$$\hat{m}_t = \hat{\beta}_0 = w'y = \sum_{j=-h}^{h} w_j y_{t-j}$$

➜ estimation par filtre symétrique.

Cas général

On montre :

$$\sum_{j=-h}^{h} w_j = 1, \quad \forall r \in [\![1, d]\!] : \sum_{j=-h}^{h} j^r w_j = 0$$

➔ le filtre w conserve les tendance polynomiales de degrés d

Cas général

On montre :

$$\sum_{j=-h}^{h} w_j = 1, \quad \forall r \in [\![1, d]\!] : \sum_{j=-h}^{h} j^r w_j = 0$$

➔ le filtre w conserve les tendance polynomiales de degrés d

Filtre Henderson $d = 3$:

$$w_j = \kappa_j \frac{(S_4 - S_2 j^2)}{S_0 S_4 - S_2^2} \text{ avec } S_r = \sum_{j=-h}^{h} \kappa_j j^r$$

κ_j trouvés pour maximiser le *smoothness* ($S = \sum(\nabla^3 w_j)^2$) :

$$\kappa_j = \left[(h+1)^2 - j^2\right]\left[(h+2)^2 - j^2\right]\left[(h+3)^2 - j^2\right]$$

Filtres asymétriques

3 approches possibles :

1. *Filtre asymétrique direct* par un polynôme local calculé sur $(y_t)_{t\in[\![n-h,n]\!]}$
2. Prévision à l'horizon h (ou backast) et utilisation d'un filtre symétrique sur $(\hat{y}_{n+j|n})_{j\in[\![1,h]\!]}$
3. Construction d'un filtre minimisant le carré des révisions sous certaines contraintes.

Filtres asymétriques

3 approches possibles :

1. *Filtre asymétrique direct* par un polynôme local calculé sur $(y_t)_{t\in[\![n-h,n]\!]}$
2. Prévision à l'horizon h (ou backast) et utilisation d'un filtre symétrique sur $(\hat{y}_{n+j|n})_{j\in[\![1,h]\!]}$
3. Construction d'un filtre minimisant le carré des révisions sous certaines contraintes.

Propriété

Si on suppose données disponibles jusqu'en $t+q$, $\hat{m}_{t+q|n}$ calculés à partir de polynômes de degré d :

$$\text{méthode 1} \iff \text{méthode 2}$$

Filtres asymétriques directs

- Filtre asymétrique très concentré sur l'observation courante : forte discontinuité entre filtre en temps réel ($q = 0$) et lorsque l'on a une observation ($q = 1$)
  ➔ filtre en temps réel sans biais si série polynôme degré d mais variance élevée

Filtres asymétriques directs

- Filtre asymétrique très concentré sur l'observation courante : forte discontinuité entre filtre en temps réel ($q = 0$) et lorsque l'on a une observation ($q = 1$)
  ➔ filtre en temps réel sans biais si série polynôme degré d mais variance élevée
- d fixé : ✚ longueur du filtre grand ($q \nearrow$), ✚ poids associé à la date courante petit
- h et q fixés : ✚ $d \nearrow$, ✚ poids associé à la date courante grand

Généralisation

On considère maintenant :

$$y = U\gamma + Z\delta + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, D)$$

Objectif : trouver filtre v qui minimise révisions sous certaines contraintes ($U_p'v = U'w$, $U = X \implies$ le filtre v reproduit les polynômes de degré d)

Équivalent à minimiser :

$$\varphi(v) = \underbrace{\underbrace{(v - w_p)'D_p(v - w_p) + w_f'D_f w_f}_{\text{variance err révision}} + \underbrace{[\delta'(Z_p'v - Z'w)]^2}_{biais^2}}_{\text{MSE révisions}} + 2l'(U_p'v - U'v)$$

L'erreur de révision est :

$$\hat{m}_{t|t} - \hat{m}_t = v'y_p - w'y = (v'Z_p - w'Z)\delta + v'\varepsilon_p - w'\varepsilon$$

Cas particuliers

1. Pour $D = K^{-1}$ et $U = X$ (pas de biais) on retrouve le filtre symétrique.

2. Lorsque le filtre symétrique w est le filtre de Henderson, $U = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$,

   $Z = \begin{pmatrix} -h \\ \vdots \\ h \end{pmatrix}$, $D = \sigma^2 I$ on retrouve les filtres de Musgrave
   $\iff$ coefficients trouvés en faisant prévision par une extrapolation linéaire
   δ^2/σ^2 est lié au ratio $R = \bar{I}/\bar{C}$.

Idée

Relâcher certaines contraintes : arbitrage biais variance. Si U sous ensemble de X : on reproduit polynômes de degré $d^* < d$

➔ poids mieux distribués, peut réduire la variance

Idée

Relâcher certaines contraintes : arbitrage biais variance. Si U sous ensemble de X : on reproduit polynômes de degré $d^* < d$
➔ poids mieux distribués, peut réduire la variance

Trois classes de filtres en approximant le filtre symétrique de Henderson :

1. *Linear-Constante* (LC) : y_t linéaire et v conserve constante ➔ dépend de δ_1^2/σ^2
2. *Quadratic-Linear* (QL) : y_t degré 2 et v conserve polynôme degré 1 ➔ dépend de δ_2^2/σ^2
3. *Cubic-Quadratic* (CQ) : y_t degré 3 et v conserve polynôme degré 2 ➔ dépend de δ_3^2/σ^2

Idée

Relâcher certaines contraintes : arbitrage biais variance. Si U sous ensemble de X : on reproduit polynômes de degré $d^* < d$
➔ poids mieux distribués, peut réduire la variance

Trois classes de filtres en approximant le filtre symétrique de Henderson :

1. *Linear-Constante* (LC) : y_t linéaire et v conserve constante ➔ dépend de δ_1^2/σ^2
2. *Quadratic-Linear* (QL) : y_t degré 2 et v conserve polynôme degré 1 ➔ dépend de δ_2^2/σ^2
3. *Cubic-Quadratic* (CQ) : y_t degré 3 et v conserve polynôme degré 2 ➔ dépend de δ_3^2/σ^2

Plus δ_i^2/σ^2 grand, plus filtre donne d'importance aux observations courantes et détectera les points de retournement (timeliness).

$\delta_3^2/\sigma^2 = 0$: filtre QL avec $\delta_2^2/\sigma^2 \to +\infty$

$\delta_3^2/\sigma^2 \to +\infty$: retrouve filtre asymétriques directs

Choix des paramètres

Dans les applications, δ_i^2/σ^2 trouvé en minimisant les révisions :

$$MSRE = \frac{1}{n-2h-1}\sum_{t=h+1}^{n-h}(\hat{m}_t - \hat{m}_{t|t})^2$$

Choix des paramètres

Dans les applications, δ_i^2/σ^2 trouvé en minimisant les révisions :

$$MSRE = \frac{1}{n-2h-1}\sum_{t=h+1}^{n-h}(\hat{m}_t - \hat{m}_{t|t})^2$$

- Avec filtre LC biais en cas de point de retournement
- CQ approche mieux les données mais a une variance plus grande : estimations parfois moins efficaces.
- QL : compromis entre flexibilité et lissage

Question en suspens

- Dans la désaisonnalisation, que se passe-t-il quand la tendance est mal estimée ? mauvaise décomposition TC+I ?
- Dans la désaisonnalisation, les filtres sont appliqués sur la séries linéarisées :
  - les points de retournement sont-ils vraiment détectables ? Ne sont-ils pas considérés comme du bruit (correction automatique de X11) ?
  - impact d'une mauvaise spécification/estimation lors du pré-ajustement ?
- on considère des tendances déterministes, quid des retournements liés aux tendances stochastiques ?