

DÉSAISONNALISER UNE SÉRIE TEMPORELLE



4 - Méthode X13-ARIMA

ALAIN QUARTIER-LA-TENTE

Objectifs de cette séquence

Cette séquence a pour objectif de vous présenter les concepts de base relatifs à la méthode X13-ARIMA.

Après cette séquence vous saurez :

- les vocabulaires relatifs à la méthode X13-ARIMA
- la structure de la méthode X13-ARIMA en deux étapes
- le fonctionnement et la finalité de la phase de pré-ajustement
- le principe de la décomposition

Questions de positionnement

Quels usages de la phase de pré-ajustement dans la méthode X13-ARIMA ?

Quels sont les Outliers les plus couramment détectés ?

À quoi sert un modèle ARIMA dans la méthode X13-ARIMA ?

Qu'est ce que X11 ?

Qu'est ce que le principe itératif dans la décomposition ?

X13-ARIMA

X pour eXperience. . .

Deux modules :

- REG-ARIMA : phase de pré-ajustement
Correction préalable par régression linéaire des points aberrants, ruptures de tendance, effets de calendrier.

Modélisation ARIMA : pour prolonger la série brute afin de résoudre partiellement le problème des fins de série lié aux moyennes mobiles symétriques.

- X11 : phase de décomposition

Décomposition de la série en tendance-cycle, saisonnalité et irrégulier, à l'aide de moyennes mobiles.

Sommaire

1. Phase de pré-ajustement : modèle Reg-ARIMA

1.1 Série linéarisée

1.2 Outliers et autres régresseurs

1.3 Modèle ARIMA

2. Phase de décomposition (X11)

3. Conclusion

Linéariser la série

Les moyennes mobiles, comme les modèles ARIMA, sont des opérateurs linéaires. Il faut donc préalablement **supprimer les « non-linéarités »** des séries (points aberrants, effets de calendrier, ruptures...)

Par régression linéaire sur les :

- outliers (points aberrants et ruptures)
- effets de calendrier ➡ séquence de demain matin
- autres régresseurs éventuels (ex : température moyenne, vacances scolaires...)

Série linéarisée = $(Y_t - \sum \hat{\alpha}_i X_{it})$ où les X_i modélisent les « non-linéarités »

Les principaux types d'outliers

Choc ponctuel

Additive outlier (AO)

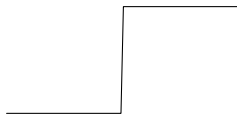
Affecte l'Irrégulier



Changement de niveau

Level Shift (LS)

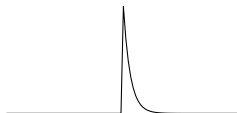
Affecte la Tendance



Changement de niveau transitoire

Transitory Change (TC)

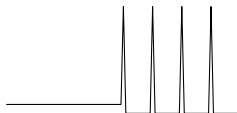
Affecte l'Irrégulier



Rupture de profil saisonnier

Seasonal Outlier (SO)

Affecte la Composante Saisonnière

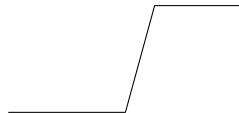


Autres outliers *atypiques*

Effet de rampe

Ramp effect (RP)

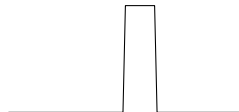
Affecte la Tendance



Rupture en niveau temporaire

Tempory Level Shift (TLS)

Affecte la Tendance



Pure seasonal outlier (PSO)

Affecte la Composante Saisonnière



Variables de régression associées aux outliers

$$AO_t^{t_0} = \begin{cases} 1 & t = t_0 \\ 0 & t \neq t_0 \end{cases}$$

$$LS_t^{t_0} = \begin{cases} -1 & t < t_0 \\ 0 & t \geq t_0 \end{cases}$$

$$TC_t^{t_0} = \begin{cases} 0 & t < t_0 \\ \alpha^{t-t_0} & t \geq t_0 \end{cases}$$

avec $\alpha \in]0, 1[$ ($\alpha = 0,7$ par défaut)

$$SO_t^{t_0} = \begin{cases} 1 & t < t_0, t \text{ même mois que } t_0 \\ -\frac{1}{s-1} & t < t_0, t \text{ autre mois que } t_0 \\ 0 & t \geq t_0 \end{cases}$$

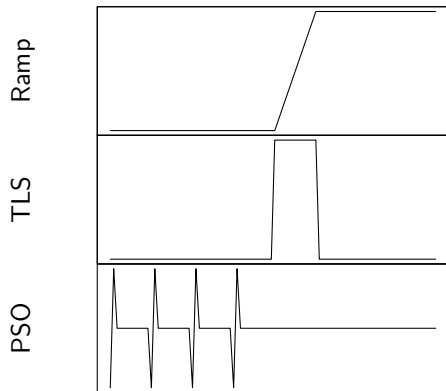
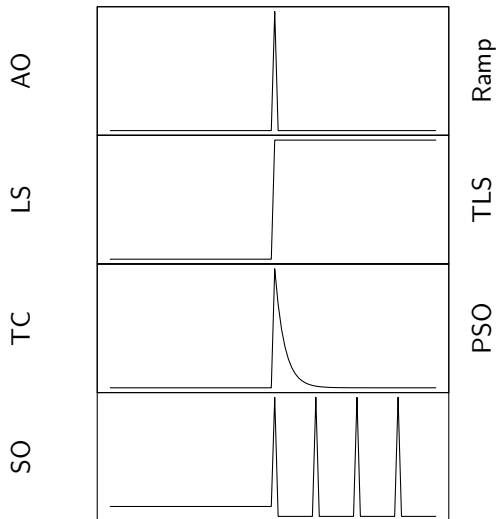
avec s la périodicité de la série

$$RP_t^{(t_0, t_1)} = \begin{cases} -1 & t \geq t_0 \\ -\frac{t-t_0}{t_1-t_0} - 1 & t_0 < t < t_1 \\ 0 & t \leq t_1 \end{cases}$$

$$TLS_t^{(t_0, t_1)} = \begin{cases} 0 & t < t_0 \\ 1 & t_0 \leq t \leq t_1 \\ 0 & t > t_1 \end{cases}$$

$$PSO_t^{(t_0, j)} = \begin{cases} 1 & t \text{ même mois que } j, t < t_0 \\ 0 & t \text{ autre mois que } j \text{ ou } t \geq t_0 \\ -1 & t \text{ même mois que } t_0, t < t_0 \end{cases}$$

Outliers : résumé graphique



La modélisation Reg-ARIMA

Le modèle Reg-ARIMA :

$$\left(Y_t - \sum \hat{\alpha}_i X_{it}\right) \sim ARIMA(p, d, q)(P, D, Q)$$

où les X_i modélisent les « non-linéarités »

Le modèle ARIMA capture toute l'information temporelle, le résidu du modèle Reg-ARIMA est un bruit blanc

Le modèle ARIMA est utilisé pour faire la prévision sur la série linéarisée

La structure du modèle Reg-ARIMA, ainsi que sa détermination seront vues en détail lors d'une autre session

Sommaire

1. Phase de pré-ajustement : modèle Reg-ARIMA

2. Phase de décomposition (X11)

2.1 Les moyennes mobiles

2.2 Le principe itératif de X11

2.3 Les étapes X11

3. Conclusion

Les moyennes mobiles (1/2)

La décomposition X11 utilise les filtres que sont les moyennes mobiles

Une MM est un opérateur linéaire :

Linéarité : $M(X_t + Y_t) = M(X_t) + M(Y_t)$

3 types de MM utilisés par X11 :

1. Moyennes mobiles simples (pour extraire la tendance) $\rightarrow M_{2 \times 12}$ (ou $M_{2 \times 4}$ pour séries trim)
 - conserve la tendance linéaire (locale) :
 $M(at + b) = at + b \implies M(TC_t) \simeq TC_t$
 - élimine la saisonnalité stable : $M(S_t) = 0$
 - réduit le bruit : $\mathbb{V}[M(I_t)] \ll \mathbb{V}[I_t] \implies M(I_t) \simeq 0$

Les moyennes mobiles (2/2)

2. Moyennes mobiles de Macurves (pour extraire la saisonnalité) $\rightarrow M_{3 \times 5}$
 - réduit le bruit
 - conserve la saisonnalité stable : $M(S_t) = S_t$
3. Moyennes mobiles de Henderson (pour extraire la tendance) $\rightarrow H_{13}$ (ou H_7 pour séries trim)
 - conserve la tendance polynomiale (ordre 3) :
 $M(at^3 + bt^2 + ct + d) = at^3 + bt^2 + ct + d$
 - n'élimine pas la saisonnalité : $M(S_t) \neq 0$ et $M(S_t) \neq S_t$
 - réduit le bruit au maximum

Principe itératif de X11 (1/2)

Une première estimation de la CVS :

1. Estimation de la **tendance-cycle** par moyenne mobile 2×12 (ou 2×4) :

$$TC_t^{(1)} = M_{2 \times 12}(X_t)$$

2. Estimation de la composante **saisonnier-irrégulier** :

$$(S_t + I_t)^{(1)} = X_t - TC_t^{(1)}$$

3. Estimation de la composante **saisonnrière** par moyenne mobile 3×3 sur **chaque mois** :

$$S_t^{(1)} = M_{3 \times 3} \left[(S_t + I_t)^{(1)} \right] \text{ et normalisation } Snorm_t^{(1)} = S_t^{(1)} - M_{2 \times 12} \left(S_t^{(1)} \right)$$

4. Estimation de la série corrigée des variations saisonnières :

$$Xsa_t^{(1)} = (TC_t + I_t)^{(1)} = X_t - Snorm_t^{(1)}$$

Principe itératif de X11 (2/2)

Une seconde estimation de la CVS :

1. Estimation de la **tendance-cycle** par moyenne de Henderson (généralement 13 termes ou 5 pour séries trim) :

$$TC_t^{(2)} = H_{13}(Xsa_t^{(1)})$$

2. Estimation de la composante **saisonnier-irrégulier** :

$$(S_t + I_t)^{(2)} = X_t - TC_t^{(2)}$$

3. Estimation de la composante **saisonnrière** par moyenne mobile 3×5 (généralement) sur **chaque mois** :

$$S_t^{(2)} = M_{3 \times 5} \left[(S_t + I_t)^{(2)} \right] \text{ et normalisation } Snorm_t^{(2)} = S_t^{(2)} - M_{2 \times 12} \left(S_t^{(2)} \right)$$

4. Estimation de la série corrigée des variations saisonnières :

$$Xsa_t^{(2)} = (TC_t + I_t)^{(2)} = X_t - Snorm_t^{(2)}$$

Les étapes de X11

7 grandes étapes A à G

Étape A : ajustements préalables (n'est plus utile)

Étapes B et C : corrections automatiques (sur l'Irrégulier)

Étape D : désaisonnalisation finale

Étapes E, F et G : indicateurs et graphiques pour juger de la qualité de la décomposition (n'est plus utile avec JDemetra+)

Les corrections automatiques de X11 (2/2)

Il existe un algorithme de détection et de correction des outliers et des effets de calendrier dans X11, **indépendant du module reg-ARIMA**.

Ces détections sont faites sur la **composante irrégulière**.

Elles s'effectuent dans les étapes B, C et D.

Quand on n'autorise pas la correction des outliers dans la phase de pré-ajustement, X11 le fait quand même dans la phase de décomposition, avec pour conséquences :

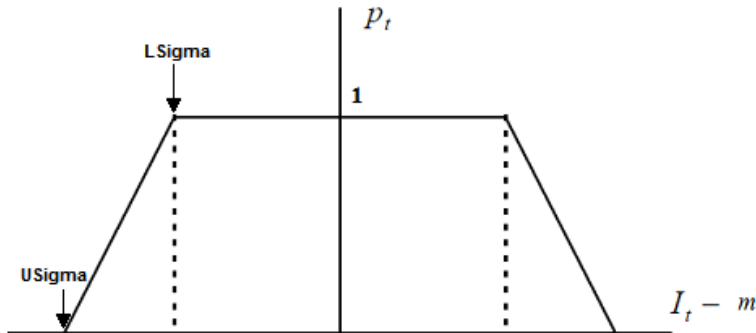
- moins de visibilité et de maîtrise sur les outliers, leur nombre et leur nature (AO, LS ou TC) ;
- la prévision du modèle Reg-ARIMA sera moins précise.

Les corrections automatiques de X11 (2/2)

Étape 1 : calcul d'un écart-type mobile de l'irrégulier sur des intervalles de 5 ans

- Tout point I_t de l'année centrale telle que $|I_t - m| > 2,5\sigma_t$ est considéré comme aberrant et « supprimé »
- On recalcule l'écart-type (estimation plus robuste)

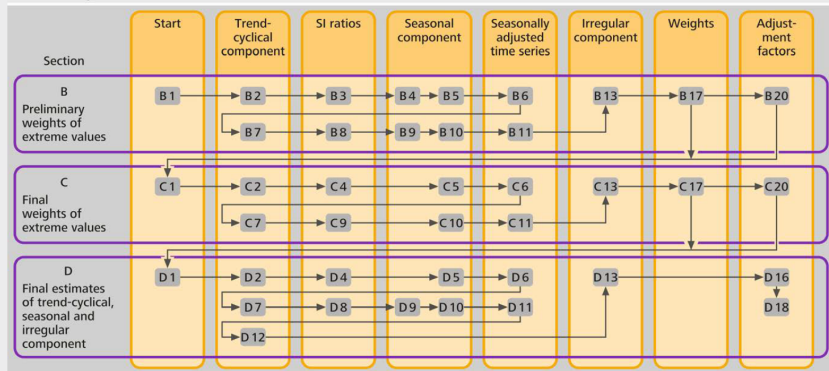
Étape 2 : On affecte un poids à chaque I_t en fonction du σ associé (l'année centrale étant l'année de t)



Bilan des étapes de X11

Basic principle of the X-11 seasonal adjustment algorithm in JDemetra+

Workflow diagram



Deutsche Bundesbank

S3PR0037B.Chart

Sommaire

1. Phase de pré-ajustement : modèle Reg-ARIMA

2. Phase de décomposition (X11)

3. Conclusion

Les essentiels

- X13-ARIMA travaille en deux phases : pré-ajustement et décomposition
- Le pré-ajustement linéarise (par régression) et prolonge les séries en faisant des prévisions (par modèle ARIMA)
- La décomposition X11 estime la composante Saisonnière
- X11 utilise plusieurs moyennes mobiles ayant des propriétés complémentaires, de manière successive et itérative : principe itératif de X11