Le Lissage exponentiel

Le lissage exponentiel simple

Le lissage exponentiel double

Le lissage de Holt

Le lissage de Winters

Le lissage Généralisé

‘‘Les modèles de Gardner’’

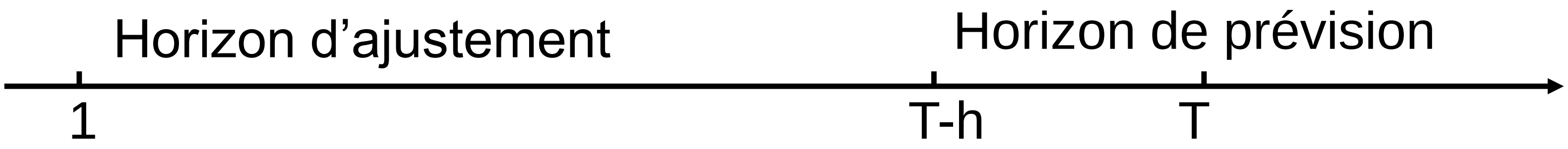
‘‘Les modèles ETS’’

Modélisation des séries temporelles

La série temporelle y_t est disponible pour T observations

La prévision se fera pour les h derniers instants

Le modèle est ajusté à la série pour les instants : t = 1, … , T-h



On peut distinguer la qualité d'ajustement et la qualité de prévision

Les paramètres du modèle sont estimés en optimisant son ajustement sur la période d'ajustement : minimiser un critère de comparaison choisi

Important : les valeurs de la série dans l'horizon de prévision ne participent pas dans le calcul des paramètres du modèle, fait dans l'horizon d'ajustement

Calcul d’un modèle de lissage exponentiel

Paramètres du modèle :

- Valeurs initiales
- Paramètres du lissages exponentiel ∈ $[0, 1]$

Choisir le critère de comparaison à minimiser pour ajuster le modèle : RMSE, MAE ou MAPE

La fonction ‘‘critère de comparaison’’ dépend des paramètres du modèle

L’estimation des paramètres se fait en calculant le minimum de la fonction critère de comparaison par application d’une procédure d’optimisation : optim

Calculer le modèle sur la période d’ajustement

Calculer les prévisions

Apprécier la qualité des prévisions

Le lissage exponentiel simple (LES)

Soit la série chronologique y_t $t = 1, \dots, T$

La série F_t du modèle ajusté a pour signification :

“prévision à partir de l’instant t pour l’instant t+1”

F_t se calcule à l’instant t

$$F_t = \alpha y_t + (1-\alpha)F_{t-1},\ t = 1, \dots, T,\ 0<\alpha<1$$

α est appelée la constante du lissage

$$F_t = \alpha y_t + (1-\alpha)F_{t-1}$$

$$F_{t-1} = \alpha y_{t-1} + (1-\alpha)F_{t-2}$$

$$\vdots$$

$$F_2 = \alpha y_2 + (1-\alpha)F_1$$

$$F_t = \sum_{j=0}^{t-2} \alpha(1-\alpha)^j y_{t-j} + (1-\alpha)^{t-1} F_1$$

On déduit alors que :

$$F_T = \sum_{j=0}^{T-2} \alpha(1-\alpha)^j y_{T-j} + (1-\alpha)^{T\text{-}1} F_1 \approx \sum_{j=0}^{T-2} \alpha(1-\alpha)^j y_{T-j} \quad \text{si T est assez grand}$$

Dans ce cas, la valeur de F_1 n'a qu'une très faible incidence sur la valeur de F_T

Le problème de minimisation suivant $\min \sum_{j=0}^{T-2} (1-\alpha)^j (y_{T-j} - a)^2$

a pour solution : $a = \dfrac{\sum_{j=0}^{T-2} \alpha(1-\alpha)^j y_{T-j}}{1-(1-\alpha)^{T-1}} \approx F_T$ si T est assez grand

L'erreur d'ajustement : $e_t = y_t - F_{t-1}$

La prévision de la série y_t à l'instant T + h est : $\hat{y}_{T+h} = F_T$

Choix optimale de la constante de lissage

On commence par choisir un des critères de comparaison, MSE, RMSE, MAE, MAPE

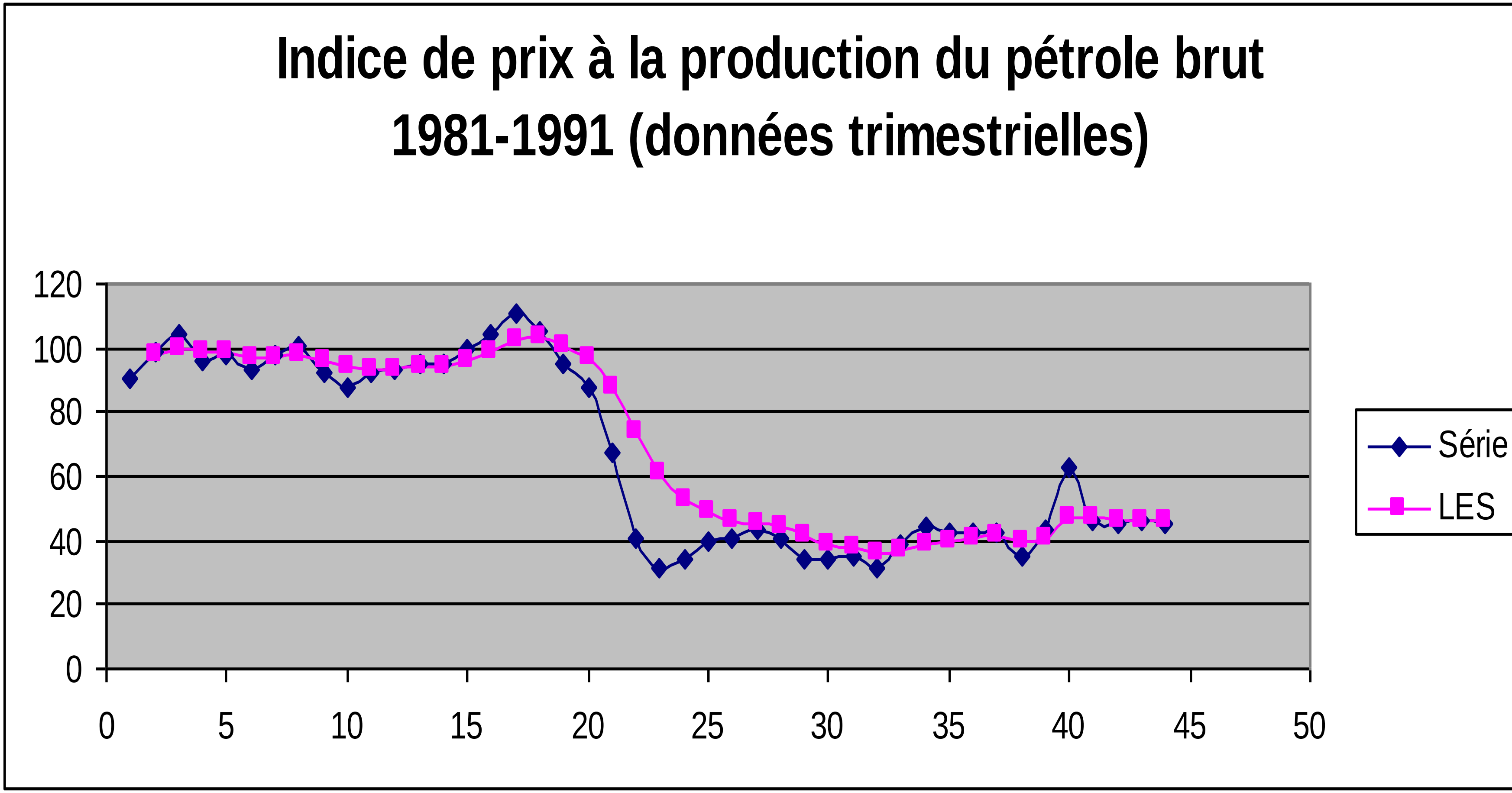
On calcule les modèles de lissage pour $\alpha = 0.1, 0.2, \ldots 0.9$

Le modèle retenu est celui qui assure le plus petit critère

Appliquons la méthode du lissage exponentiel à la série trimestrielle de l'indice du prix à la production du pétrole brute.

$\alpha = 0.3$ et F_2 est égale à la moyenne des trois premières valeurs.

On aurait pu choisir $F_1 = y_1$



1) Le lissage réalise un écrêtage de la série

2) Le lissage s'adapte au changement de niveau

Choix optimale de la constante de lissage

ALPHA	0,2	0,3	0,4	0,6	0,7	0,9
MSE	250,249	177,795	138,463	96,570	75,563	69,063
MAE	9,949	8,227	7,232	6,393	5,932	5,806
MAPE	0,221	0,173	0,146	0,122	0,110	0,106

On remarque que les critères sont minimales pour $\alpha = 0.9$

Equations du modèle du lissage exponentiel simple

Les paramètres du modèles sont : F_2 et α

On choisit le MAPE comme critère à minimiser pour estimer le modèle, c'est une fonction dans ce cas à 2 variables

Initialisation des paramètres : $F_2 = x[1]$ et $\alpha = x[2]$

Mape = 0
Pour i = 3 à T-h

$$F_i = \alpha y_i + (1-\alpha)F_{i-1}$$
$$Mape = Mape + abs(y_i - F_{i-1}) * 100/y_i$$

Fin_pour

$$Mape = Mape/(T-h-2)$$

Optimisation de la fonction Mape par la procédure optim

Calcul du modèle sur l'horizon d'ajustement : 1,…, T-h

Calcul des prévisions pour les h derniers instants :

$\hat{y}_t = F_{T-h}$, pour $t \geq T-h+1$

Le lissage exponentiel double (LED)

C'est une méthode plus générale que le LES

Elle s'adapte aux séries présentant une tendance

Pour une série y_t t = 1, … ,T, LED suggère une prévision de la forme :

$$\hat{y}_T(h) = S_T + hT_T$$

$\hat{y}_T(h)$: valeur prévue de la série à partir de T pour l'instant T+h

S_T et T_T sont choisis de façon à minimiser :

$$Q(S_T, T_T) = \sum_{j=0}^{T-1} (1-\alpha)^j (y_{T-j} - \hat{y}_{T-j}(T))^2 = \sum_{j=0}^{T-1} (1-\alpha)^j (y_{T-j} - S_T + jT_T)^2$$

Pour cela on a : $\dfrac{\partial Q(S_T, T_T)}{\partial S_T} = \dfrac{\partial Q(S_T, T_T)}{\partial T_T} = 0$

D'où : $\sum_{j=0}^{T-1}(1-\alpha)^j(y_{T-j}-S_T+jT_T)=0 \quad \sum_{j=0}^{T-1}(1-\alpha)^j j(y_{T-j}-S_T+jT_T)=0$

Si T est assez grand, remplaçons :

$\sum_{j=0}^{T-1}\beta^j$ par $\frac{1}{1-\beta}$ $\qquad \sum_{j=0}^{T-1}j\beta^j$ par $\frac{\beta}{(1-\beta)^2}$ $\qquad \sum_{j=0}^{T-1}j^2\beta^j$ par $\frac{\beta(1+\beta)}{(1-\beta)^3}$

Dans ce cas :

$$\alpha\sum_{j=0}^{T-1}(1-\alpha)^j y_{T-j}-S_T+(1-\alpha)\frac{T_T}{\alpha}=0$$

$$\alpha^2\sum_{j=0}^{T-1}(1-\alpha)^j jy_{T-j}-(1-\alpha)S_T+(1-\alpha)(2-\alpha)\frac{T_T}{\alpha}=0$$

On pose : $E_1(T) = \alpha \sum_{j=0}^{T-1} (1-\alpha)^j y_{T-j}$ $E_2(T) = \alpha^2 \sum_{j=0}^{T-1} (j+1)(1-\alpha)^j y_{T-j}$

D’où

$$E_1(T) - S_T + (1-\alpha)\frac{T_T}{\alpha} = 0$$

$$E_2(T) - \alpha E_1(T) - (1-\alpha)S_T + (1-\alpha)(2-\alpha)\frac{T_T}{\alpha} = 0$$

Par suite :

$$E_1(T) = S_T - (1-\alpha)\frac{T_T}{\alpha}$$

$$E_2(T) = \alpha E_1(T) + (1-\alpha)S_T - (1-\alpha)(2-\alpha)\frac{T_T}{\alpha} = 0$$

$$E_2(T) = \alpha S(T) - (1-\alpha)T_T + (1-\alpha)S_T - (1-\alpha)(2-\alpha)\frac{T_T}{\alpha} = 0$$

$$E_2(T) = S_T - 2(1-\alpha)\frac{T_T}{\alpha}$$

D'où $T_T = \frac{\alpha}{1-\alpha}\left(E_1(T) - E_2(T)\right)$ $\qquad S_T = 2E_1(T) - E_2(T)$

Formule de mise à jour :

$$E_1(T) = \alpha\sum_{j=0}^{T-1}(1-\alpha)^j y_{T-j} = \alpha y_T + \alpha\sum_{i=0}^{T-2}(1-\alpha)^{i+1} y_{T-i-1} = \alpha y_T + (1-\alpha)E_1(T-1)$$

$$E_2(T) = \alpha^2\sum_{j=0}^{T-1}(j+1)(1-\alpha)^j y_{T-j} = \alpha^2\sum_{j=0}^{T-1} j(1-\alpha)^j y_{T-j} + \alpha E_1(T)$$

$$E_2(T) = \alpha^2\sum_{j=1}^{T-1} j(1-\alpha)^j y_{T-j} + \alpha E_1(T) = \alpha^2\sum_{i=0}^{T-2}(i+1)(1-\alpha)^{i+1} y_{T-i-1} + \alpha E_1(T)$$

$$E_2(T) = (1-\alpha)E_2(T-1) + \alpha^2 y_T + \alpha(1-\alpha)E_1(T-1)$$

A partir des formules de mise à jour de $E_1(T)$ et $E_2(T)$, il est possible de déduire les formules de mise à jour de S_T et T_T

Mise à jour de S_T

$$S_T = 2E_1(T) - E_2(T)$$

$$S_T = 2\alpha y_T + 2(1-\alpha)E_1(T-1) - (1-\alpha)E_2(T-1) - \alpha^2 y_T - \alpha(1-\alpha)E_1(T-1)$$

$$S_T = (2\alpha - \alpha^2)y_T + (1-\alpha)E_1(T-1)(2-\alpha) - (1-\alpha)E_2(T-1)$$

$$S_T = (2\alpha - \alpha^2)y_T + (1-\alpha)(2-\alpha)(S_{T-1} - (1-\alpha)\frac{T_{T-1}}{\alpha}) - (1-\alpha)(S_{T-1} - 2(1-\alpha)\frac{T_{T-1}}{\alpha})$$

$$S_T = (2\alpha - \alpha^2)y_T + (1-\alpha)S_{T-1}(2-\alpha-1) + (1-\alpha)^2\frac{T_{T-1}}{\alpha}(-2+\alpha+2)$$

$$S_T = (2\alpha - \alpha^2)y_T + (1-\alpha)^2(S_{T-1} + T_{T-1})$$

$$S_T = \lambda y_T + (1-\lambda)(S_{T-1} + T_{T-1})$$ avec $\lambda = 2\alpha - \alpha^2$

$$E_1(T) = S_T - (1-\alpha)\frac{T_T}{\alpha} \qquad E_2(T) = S_T - 2(1-\alpha)\frac{T_T}{\alpha}$$

Mise à jour de T_T

$$T_T = \frac{\alpha}{1-\alpha}\left(E_1(T) - E_2(T)\right)$$

$$T_T = \frac{\alpha}{1-\alpha}\left(\alpha y_T + (1-\alpha)E_1(T-1) - (1-\alpha)E_2(T-1) - \alpha^2 y_T - \alpha(1-\alpha)E_1(T-1)\right)$$

$$T_T = \frac{\alpha}{1-\alpha}\left(\alpha(1-\alpha)y_T + (1-\alpha)^2 E_1(T-1) - (1-\alpha)E_2(T-1)\right)$$

$$T_T = \alpha^2 y_T + \alpha(1-\alpha)E_1(T-1) - \alpha E_2(T-1)$$

$$T_T = \alpha^2 y_T + \alpha(1-\alpha)S_{T-1} - (1-\alpha)^2 T_{T-1} - \alpha S_{T-1} + 2(1-\alpha)T_{T-1}$$

$$T_T = \alpha^2 y_T - \alpha^2 S_{T-1} + (1-\alpha^2)T_{T-1}$$

$$E_1(T) = S_T - (1-\alpha)\frac{T_T}{\alpha} \qquad E_2(T) = S_T - 2(1-\alpha)\frac{T_T}{\alpha}$$

Autre formulation de la mise à jour de T_T

$$y_T = \frac{1}{\lambda}\left(S_T - (1-\lambda)(S_{T-1} + T_{T-1})\right) = \frac{1}{2\alpha - \alpha^2}\left(S_T - (1-\alpha)^2(S_{T-1} + T_{T-1})\right)$$

$$T_T = \frac{\alpha^2}{2\alpha - \alpha^2}\Big(S_T - (1-\alpha)^2 (S_{T\text{-}1} + T_{T\text{-}1})\Big) - \alpha^2 S_{T-1} + (1-\alpha^2) T_{T-1}$$

$$T_T = \frac{\alpha}{2-\alpha}\Big(S_T - (1-\alpha)^2 (S_{T\text{-}1} + T_{T\text{-}1})\Big) - \alpha^2 S_{T-1} + (1-\alpha^2) T_{T-1}$$

$$T_T = \frac{\alpha}{2-\alpha} S_T - \left(\frac{\alpha(1-\alpha)^2}{2-\alpha} + \alpha^2\right) S_{T-1} + \left(1 - \alpha^2 - \frac{\alpha(1-\alpha)^2}{2-\alpha}\right) T_{T-1}$$

$$T_T = \frac{\alpha}{2-\alpha} S_T - \frac{\alpha}{2-\alpha} S_{T-1} + \left(1 - \frac{\alpha}{2-\alpha}\right) T_{T-1}$$

$$T_T = \mu\left(S_T - S_{T-1}\right) + \left(1-\mu\right) T_{T-1},\ \text{avec } \mu = \frac{\alpha}{2-\alpha}$$

$$S_t = \lambda y_t + (1-\lambda)(S_{t-1} + T_{t-1}) \quad \text{avec } \lambda = 2\alpha - \alpha^2$$

$$T_t = \mu(S_t - S_{t-1}) + (1-\mu)T_{t-1}, \quad \text{avec } \mu = \frac{\alpha}{2-\alpha}$$

La série ajustée : $F_{t-1} = S_{t-1} + T_{t-1}$

L'erreur d'ajustement : $e_t = y_t - F_{t-1}$

$$\alpha = 0.3, S_2 = \frac{y_1 + y_2 + y_3}{3}, T_2 = \frac{y_3 - y_1}{2} \quad \leftarrow \textbf{Initialisation}$$

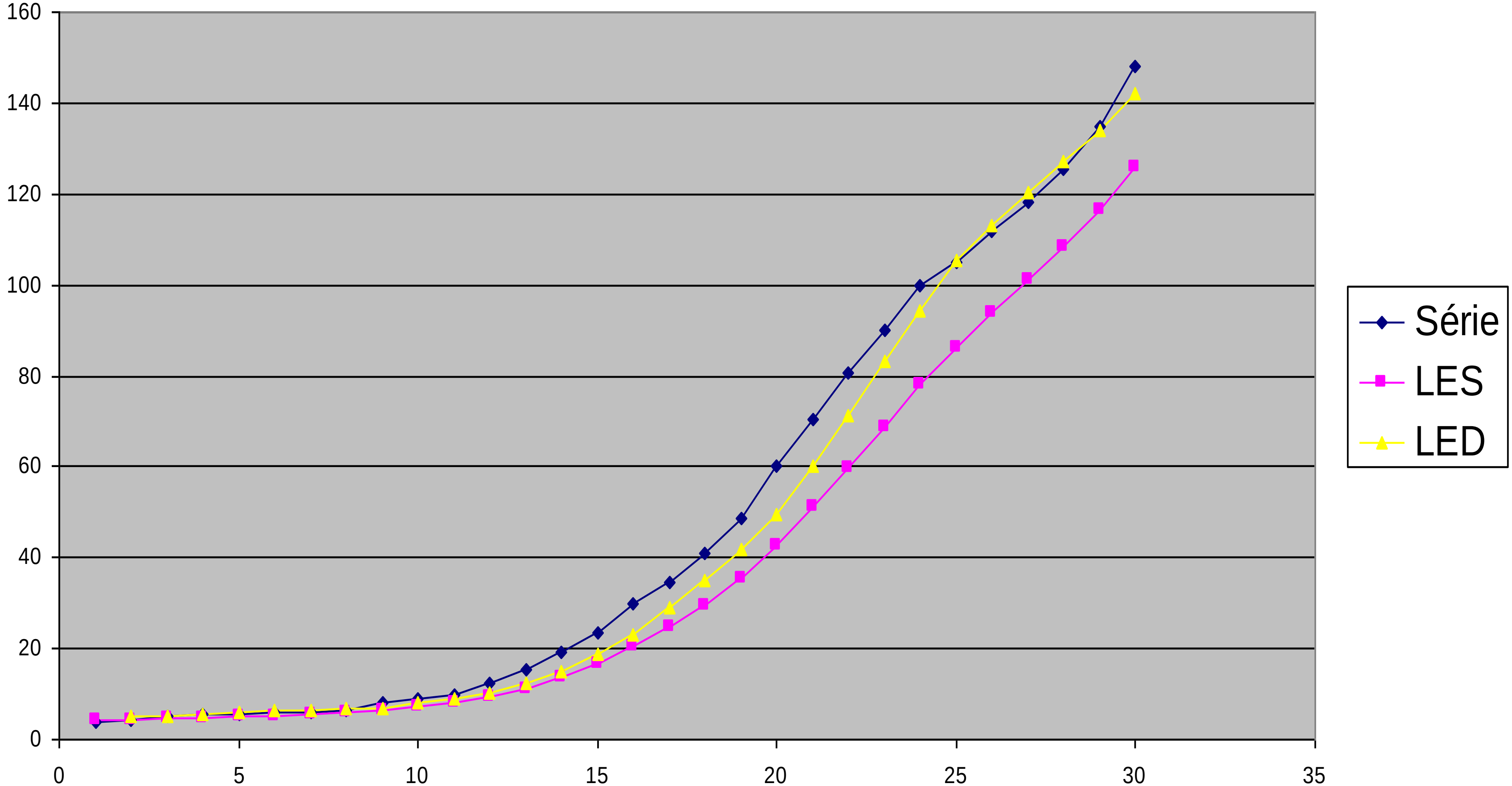
Autre possibilité de calculs des valeurs initiales : $S_1 = y_1$ et $T_1 = 0$

Remarque :

Dés que le nombre d'observations est supérieur à une dizaine, les valeurs lissées en fin de période sont très peu sensibles au choix des valeurs initiales

Appliquons le LED à la série "indice de salaire annuel moyen de l'industrie italienne 1962 - 1991"

Indice de salaire annuel moyen de l'industrie italienne 1962-1991



Le rattrapage de la pente est autant plus rapide que le coefficient α est élevé

Choix optimale de la constante de lissage

On calcule le LED pour différentes valeurs de α

alpha	0,3	0,2	0,4	0,6	0,8	0,9
mse	23,07	70,09	11,24	4,974	3,316	3,059
mae	3,528	6,245	2,63	1,564	1,143	1,125
mape	9,86	14,49	7,437	4,623	3,321	3,257

La valeur optimale de α est 0.9

Equations du modèle du lissage exponentiel double

Les paramètres du modèles sont : S_2, T_2 et α

On choisit le MAPE comme critère à minimiser pour estimer le modèle, c'est une fonction dans ce cas à 3 variables

Initialisation des paramètres : $S_2 = x[1], T_2 = x[2]$ et $\alpha = x[3]$

λ = 2α – α² et $\mu = \frac{\alpha}{2-\alpha}$

Mape = 0
Pour i = 3 à T-h

$$S_i = \lambda y_i + (1-\lambda)(S_{i-1} + T_{i-1})$$
$$T_i = \mu(S_i - S_{i-1}) + (1-\mu)T_{i-1}$$
$$Mape = Mape + abs(y_i - S_{i-1} - T_{i-1}) * 100/y_i$$

Fin_pour

$$Mape = Mape/(T - h - 2)$$

Optimisation de la fonction Mape par la procédure optim

Calcul du modèle sur l'horizon d'ajustement : 1,…, T-h

Calcul des prévisions pour les h derniers instants :
Pour i = 1 à h

$$\hat{y}_{T-h+i} = S_{T-h} + iT_{T-h}$$

Fin_pour

Le lissage de Holt (séries non saisonnière)

Le lissage de Holt s'applique aux séries chronologiques sans composantes saisonnières et à tendance localement linéaire.

Mise à jour du niveau $S_T = \lambda y_T + (1-\lambda)(S_{T\text{-}1} + T_{T\text{-}1})$

Mise à jour de la pente $T_T = \mu\left(S_T - S_{T-1}\right) + \left(1-\mu\right)T_{T-1}$

Contrairement au LED, λ et μ sont des paramètres quelconques compris entre 0 et 1.

La prévision à la date T, pour l'horizon h est : $\hat{y}_T\left(h\right) = S_T + hT_T$

Le choix de λ et μ se fait de façon à minimiser un des critères de comparaison : MSE, MAE, MAPE, (calcul long)

Exemple : si λ, μ = 0.1, …, 0.9, dans ce cas, il y a 81 cas à traiter

Appliquons le lissage de Holt à la série ‘‘indice de salaire annuel moyen de l’industrie italienne 1962 - 1991’’

$$\lambda = 0.4, \mu = 0.3, S_2 = \frac{y_1 + y_2 + y_3}{3}, T_2 = \frac{y_3 - y_1}{2}$$

Equations du modèle du lissage exponentiel de Holt

Les paramètres du modèles sont : S_2, T_2, λ et μ

On choisit le MAPE comme critère à minimiser pour estimer le modèle, c’est une fonction dans ce cas à 4 variables

Initialisation des paramètres : $S_2 = x[1], T_2 = x[2], \lambda = x[3]$ et $\mu = x[4]$

Mape = 0
Pour i = 3 à T-h

$$S_i = \lambda y_i + (1-\lambda)(S_{i-1} + T_{i-1})$$
$$T_i = \mu(S_i - S_{i-1}) + (1-\mu)T_{i-1}$$
$$Mape = Mape + abs(y_i - S_{i-1} - T_{i-1}) * 100/y_i$$

Fin_pour

$$Mape = Mape/(T-h-2)$$

Optimisation de la fonction Mape par la procédure optim

Calcul du modèle sur l'horizon d'ajustement : 1,…, T-h

Calcul des prévisions pour les h derniers instants :
Pour i = 1 à h

$$\hat{y}_{T-h+i} = S_{T-h} + iT_{T-h}$$

Fin_pour

Le lissage de Holt-Winters (séries saisonnière)

C'est une généralisation de la méthode de Holt, où la tendance est supposée localement linéaire.

La composante saisonnière I_t, de période s (s = 4 ou 12) est introduite au modèle, d'une manière multiplicative ou additive

Les paramètres du modèle sont :

S_t : Le niveau

T_t : La pente

I_t : La composante saisonnière

s : La période de la composante saisonnière

(s = 12 données mensuelles, s = 4 données trimestrielles)

Les équations de mise à jour du **modèle additif** :

$S_t = \alpha(y_t – I_{t-s}) + (1 – \alpha)(S_{t-1} + T_{t-1})$

$T_t = \gamma(S_t – S_{t-1}) + (1 – \gamma)T_{t-1}$

$I_t = \delta(y_t – S_t) + (1 – \delta)I_{t-s}$

Les paramètres α, γ et δ sont compris entre 0 et 1.

La prévision à la date T, pour l'horizon h est : $\hat{y}_T(h) = S_T + hT_T + \hat{I}_{T+h}$

On utilise la périodicité de la composante saisonnière pour estimer $\hat{I}_{T+h}$:

$\hat{I}_{T+h} = I_{T+h-s}$ si $0 < h \leq s$

$\hat{I}_{T+h} = I_{T+h-2s}$ si $s < h \leq 2s$, et ainsi de suite …

Calcul des valeurs initiales

Les valeurs initiales des composantes saisonnières sont obtenues par la méthode de décomposition saisonnière

$$S_2 = \frac{(y_3 - I_3) + (y_2 - I_2) + (y_1 - I_1)}{3}, \; T_2 = \frac{(y_3 - I_3) - (y_1 - I_1)}{2}$$

Autre façon de calculer les valeurs initiales, pour une série trimestrielle :

Les niveaux des derniers trimestres de la 1ère année sont calculés en utilisant les MMC(4)

$T_4 = S_4 – S_3$

$I_4 = y_4 – S_4$, $I_3 = y_3 – S_3$, $I_2 = y_2 – (S_3 – T_4)$, $I_1 = y_1 – (S_3 – 2T_4)$

Initialisation des composantes saisonnières :

- Calculer les moyennes mobiles centrées sur une année donnée
- Déduire les composantes saisonnières en retranchant les moyennes mobiles centrées des valeurs de la série

La série ajustée : $F_{t-1} = S_{t-1} + T_{t-1} + I_{t-s}$

L'erreur d'ajustement : $e_t = y_t - F_{t-1}$

Remarque : l'ajustement ne peut débuter qu'à partir de la $2^{ème}$ année.

Application du lissage de Holt-Winters à la série
''Consommation trimestrielle d'essence d'aviation ''

Les équations de mise à jour du modèle multiplicatif :

$$S_t = \alpha \frac{y_t}{I_{t-s}} + (1-\alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1}$$

$$I_t = \delta \frac{y_t}{S_t} + (1-\delta)I_{t-s}$$

Les paramètres α, γ et δ sont compris entre 0 et 1.

La prévision à la date T, pour l'horizon h est : $\hat{y}_T(h) = (S_T + hT_T)\hat{I}_{T+h}$

On utilise la périodicité de la composante saisonnière pour estimer I_{T+h} :

$\hat{I}_{T+h} = I_{T+h-s}$ si $0 < h \leq s$

$\hat{I}_{T+h} = I_{T+h-2s}$ si $s < h \leq 2s$, et ainsi de suite …

Calcul des valeurs initiales

Les valeurs initiales (t = 2) des composantes saisonnière sont obtenus par la méthode de décomposition saisonnière

$$S_2 = \frac{1}{3}\left(\frac{y_3}{I_3} + \frac{y_2}{I_2} + \frac{y_1}{I_1}\right), \; T_2 = \frac{1}{2}\left(\frac{y_3}{I_3} - \frac{y_1}{I_1}\right)$$

Autre façon de calculer les valeurs initiales, pour une série trimestrielle :

Les niveaux des derniers trimestres de la 1ère année sont calculés en utilisant les MMC(4)

$T_4 = S_4 - S_3$

$$I_4 = \frac{y_4}{S_4}, I_3 = \frac{y_3}{S_3}, I_2 = \frac{y_2}{S_3 - T_4}, \; I_1 = \frac{y_1}{S_3 - 2T_4}$$

Initialisation des composantes saisonnières :

- Calculer les moyennes mobiles centrées sur une année donnée
- Déduire les composantes saisonnières en faisant le rapport des valeurs de la série avec les moyennes mobiles centrées

La série ajustée : $F_{t-1} = (S_{t-1} + T_{t-1})I_{t-s}$

L'erreur d'ajustement : $e_t = y_t - F_{t-1}$

Remarque : l'ajustement ne peut débuter qu'à partir de la 2ème année.

Application à la série : ventes des bouteilles de champagne

Le lissage généralisé

La série y_t t = 1 … T, est ajustée au voisinage de T par la fonction :

φ(t-T) = f(t-T)'a, *f(t)* et *a* sont deux vecteurs de taille *(n,1)*

a(T) est estimé en minimisant $\sum_{j=0}^{T-1} \beta^j \left(Y_{T-j} - f(-j)'a\right)^2$ = *(y - Fa)'Ω(y - Fa)*

,

$$y = \begin{pmatrix} Y_T \\ Y_{T-1} \\ \vdots \\ Y_1 \end{pmatrix} \qquad F = \begin{pmatrix} f(0)' \\ f(-1)' \\ \vdots \\ f(-T+1)' \end{pmatrix}$$

, une matrice de taille *(T,n)*

$$\Omega = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \beta^j & & \\ & & & \ddots & \\ & & & & \beta^{T-1} \end{pmatrix}$$

, β est le paramètre de lissage, 0 < β < 1

Choix de la fonction $f(t)$?

$f(t)$ peut être choisie de telle sorte que l'ajustement de la série soit : constant, linéaire, polynomiale, sinusoïdale ou exponentiel.

Pour toutes ces formes, $f(t)$ vérifie la propriété de matrice à transition fixe :

$$f(t) = Af(t-1), \text{ A matrice de taille } (n,n).$$

Ajustement constant : $\varphi(t) = a$

Dans ce cas $f(t) = 1$, le lissage généralisé coïncide avec le lissage simple et $A = 1$.

Ajustement linéaire $\varphi(t) = a_1 + a_2 t$

$$f(t) = \begin{bmatrix} 1 \\ t \end{bmatrix}, \mathrm{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \text{puisque} \begin{bmatrix} 1 \\ t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ t-1 \end{bmatrix}$$

$\varphi(t)$ est un polynôme de degré m

$f_1(t)=1$, $f_2(t) = t$, $f_3(t) = t(t-1)/2$, … , $f_{m+1}(t) = t(t-1) \ldots (t-m+1)/m!$

On a la relation $f_k(t) = f_{k-1}(t-1) + f_k(t-1)$, $k > 1$

$$f(t) = \begin{bmatrix} f_1(t) \\ \vdots \\ f_{m+1}(t) \end{bmatrix}$$

est à matrice de transition fixe si : $f(t) = Af(t-1)$

$$A = \begin{bmatrix} 1 & 0 & & \cdots & & 0 \\ 1 & 1 & & & & \\ 0 & & & \ddots & & \vdots \\ & & \ddots & \ddots & & \\ & & & & & 0 \\ 0 & & & 0 & 1 & 1 \end{bmatrix}$$

La $1^{\text{ère}}$ et la $2^{\text{ème}}$ diagonale = 1

Ajustement sinusoïdale $\varphi(t) = a_1\sin(wt) + a_2\cos(wt)$

$$f(t) = \begin{bmatrix} \sin(wt) \\ \cos(wt) \end{bmatrix}, \quad \mathrm{A} = \begin{bmatrix} \cos w & \sin w \\ -\sin w & \cos w \end{bmatrix},$$ on a f(t) = Af(t-1)

Ajustement exponentiel $\varphi(t) = ae^{\alpha t}$

$f(t) = e^{\alpha t}$, $A = e^{\alpha}$, on a f(t) = Af(t-1)

La solution du problème de minimisation :

$$\sum_{j=0}^{T-1} \beta^j \left(Y_{T-j} - f(-j)'a\right)^2 = (y - Fa)'\Omega(y - Fa)$$

$a(T) = (F'\Omega F)^{-1}F'\Omega y$, avec $F'\Omega F = \sum_{j=0}^{T-1} \beta^j f(-j) f(-j)'$ et $F'\Omega y = \sum_{j=0}^{T-1} \beta^j f(-j) Y_{T-j}$

Comme $0 < \beta < 1$, $F'\Omega F \xrightarrow[T \to +\infty]{} M$

La convergence de *F'ΩF*, nécessite $\beta < e^{2\alpha}$, pour l'ajustement exponentiel et est assurée pour les autres ajustements.

Donc â(T) = $M^{-1}F'\Omega y$, avec $F'\Omega y = \sum_{j=0}^{T-1} \beta^j Y_{T-j} f(-j)$

Equation de mise à jour de â(T) = $M^{-1}Z(T)$

$$Z(T) = \sum_{j=0}^{T-1} \beta^j Y_{T-j} f(-j) = Y_T f(0) + \sum_{j=1}^{T-1} \beta^j Y_{T-j} f(-j) \qquad j = i + 1$$

$$Z(T) = Y_T f(0) + \beta \sum_{i=0}^{T-2} \beta^i Y_{T-i-1} f(-i-1)$$

$$Z(T) = Y_T f(0) + \beta A^{-1} \sum_{i=0}^{T-2} \beta^i Y_{T-1-i} f(-i)$$

$$Z(T) = Y_T f(0) + \beta A^{-1} Z(T-1)$$

$$\hat{a}(T) = M^{-1} Y_T f(0) + M^{-1} \beta A^{-1} Z(T-1)$$

$$\hat{a}(T) = M^{-1} Y_T f(0) + M^{-1} \beta A^{-1} M \hat{a}(T-1) = g Y_T + G \hat{a}(T-1)$$

avec $g = M^{-1}f(0)$ et $G = M^{-1}\beta A^{-1}M$

La prévision de Y_{T+h} est $f(h)'\hat{a}(T)$

Expression de la matrice M dans le cas d'un ajustement polynomial de degré 2

f(j)' = [1 j j(j-1)/2]

$$F'\Omega F = \sum_{j=0}^{T-1} \beta^{j} f(-j) f(-j)' \xrightarrow[j\to+\infty]{} M$$

$$M = \begin{pmatrix} \dfrac{1}{1-\beta} & \dfrac{-\beta}{(1-\beta)^2} & \dfrac{\beta}{(1-\beta)^3} \\ \dfrac{-\beta}{(1-\beta)^2} & \dfrac{\beta(1+\beta)}{(1-\beta)^3} & \dfrac{-2\beta^2-\beta}{(1-\beta)^4} \\ \dfrac{\beta}{(1-\beta)^3} & \dfrac{-2\beta^2-\beta}{(1-\beta)^4} & \dfrac{\beta^3+4\beta^2+\beta}{(1-\beta)^5} \end{pmatrix}$$

Modèle de lissage généralisé pour ajuster une série mensuelle

$$f(j) = \begin{bmatrix} 1 \\ j \\ j(j-1)/2 \\ \sin(wj) \\ \cos(wj) \\ \sin(2wj) \\ \cos(2wj) \\ \sin(3wj) \\ \cos(3wj) \\ \sin(4wj) \\ \cos(4wj) \\ \sin(5wj) \\ \cos(5wj) \\ \cos(6wj) \\ \end{bmatrix}$$

La matrice M est une matrice carrée d'ordre 14

w = $2\pi/12$

Calcul de la matrice A

$$A=\begin{bmatrix} 1 & 0 & & & & & & & & & & \cdots & & & 0 \\ 1 & 1 & & & & & & & & & & & & & \\ 0 & 1 & 1 & & & & & & & & & & & & \\ & & & a & b & & & & & & & & & & \\ & & & -b & a & & & & & & & & & & \\ & & & & & a & b & & & & & & & & \\ & & & & & -b & a & & & & & & & & \vdots \\ \vdots & & & & & & & a & b & & & & & & \\ & & & & & & & -b & a & & & & & & \\ & & & & & & & & & a & b & & & & \\ & & & & & & & & & -b & a & & & & \\ & & & & & & & & & & & a & b & & \\ & & & & & & & & & & & -b & a & & 0 \\ 0 & & & & & & & \cdots & & & & & 0 & & -1 \end{bmatrix}$$

a = cos(w) et b = sin(w)

La matrice A vérifie :

$$A\big(f(j)\big) = f(j+1), j = 1,2, \ldots$$

$$F'\Omega F = \sum_{j=0}^{T-1} \beta^j f(-j) f(-j)' \xrightarrow[j \to +\infty]{} M$$

Les modèles de Gardner

Tableau de sélection du modèle de prévision

Cas	Série ayant la plus petite variance	Modèle
A	X_t	Lissage simple
B	$(1-B)X_t$	DA-N
C	$(1-B)^2X_t$	Lissage de Holt
D	$(1-B^s)X_t$	N-M
E	$(1-B)(1-B^s)X_t$	DA-M
F	$(1-B^2)(1-B^s)X_t$	H-W multiplicatif

$BX_t = X_{t-1}$, $B^2X_t = X_{t-2}$, $B^sX_t = X_{t-s}$

s : La période de la composante saisonnière

(s = 12 données mensuelles, s = 4 données trimestrielles)

$$(1-B)X_t = X_t - X_{t-1}$$

$$(1-B)^2X_t = X_t - 2X_{t-1} + X_{t-2}$$

$$(1-B^s)X_t = X_t - X_{t-s}$$

$$(1-B)(1-B^s)X_t = (1-B^s-B+B^{s+1})X_t = X_t - X_{t-s} - X_{t-1} + X_{t-s-1}$$

$$(1-B^2)(1-B^s)X_t = (1-B^s-B^2+B^{s+2})X_t = X_t - X_{t-s} - X_{t-2} + X_{t-s-2}$$

Le modèle DA-N : damped additive none

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + \varphi T_{t-1})$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\varphi T_{t-1}$$

La série ajustée : $F_{t-1} = S_{t-1} + \varphi T_{t-1}$

$$\hat{X}_t(h) = S_t + \sum_{i=1}^{h} \varphi^i T_t$$

Le modèle N-M : none trend multiplicative

$$S_t = \alpha \frac{X_t}{I_{t-s}} + (1-\alpha)S_{t-1}$$

La série ajustée : $F_{t-1} = S_{t-1} I_{t-s}$

$$I_t = \delta \frac{X_t}{S_t} + (1-\delta)I_{t-s}$$

$$\hat{X}_t(h) = S_t I_{t+h\text{-}s}$$

Le modèle DA-M : damped additive multiplicative

$$S_t = \alpha X_t/I_{t-s} + (1-\alpha)(S_{t-1} + \varphi T_{t-1})$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\varphi T_{t-1}$$

$$I_t = \delta(X_t/S_t) + (1-\delta)I_{t-s}$$

$$\hat{X}_t(h) = \left(S_t + \sum_{i=1}^{h} \varphi^i T_t\right) I_{t+h-s}$$

La série ajustée : $F_{t-1} = (S_{t-1} + \varphi T_{t-1})I_{t-s}$

Pour tous ces modèles, l'erreur d'ajustement est définie par :
$e_t = X_t - \hat{X}_{t-1}(1) = X_t - F_{t-1}$

Formules de prévisions à l'instant h, d'une série ajustée jusqu'à l'horizon th

Modèle	Formule de prévision
DA-N	$\hat{X}_{th+h} = \hat{X}_{th}(h) = S_{th} + \sum_{i=1}^{h} \varphi^i T_{th}$
N-M	$\hat{X}_{th+h} = \hat{\mathrm{X}}_{\mathrm{th}}(\mathrm{h}) = \mathrm{S}_{\mathrm{th}} \mathrm{I}_{\mathrm{th+h-s}}$
DA-M	$\hat{X}_{th+h} = \hat{X}_{th}(h) = \left(S_{th} + \sum_{i=1}^{h} \varphi^i T_{th} \right) I_{th+h-s}$

Longueurs T des séries chronologiques pour minimiser l’effet des valeurs initiales :

Séries trimestrielles ≥ 12*4 = 48

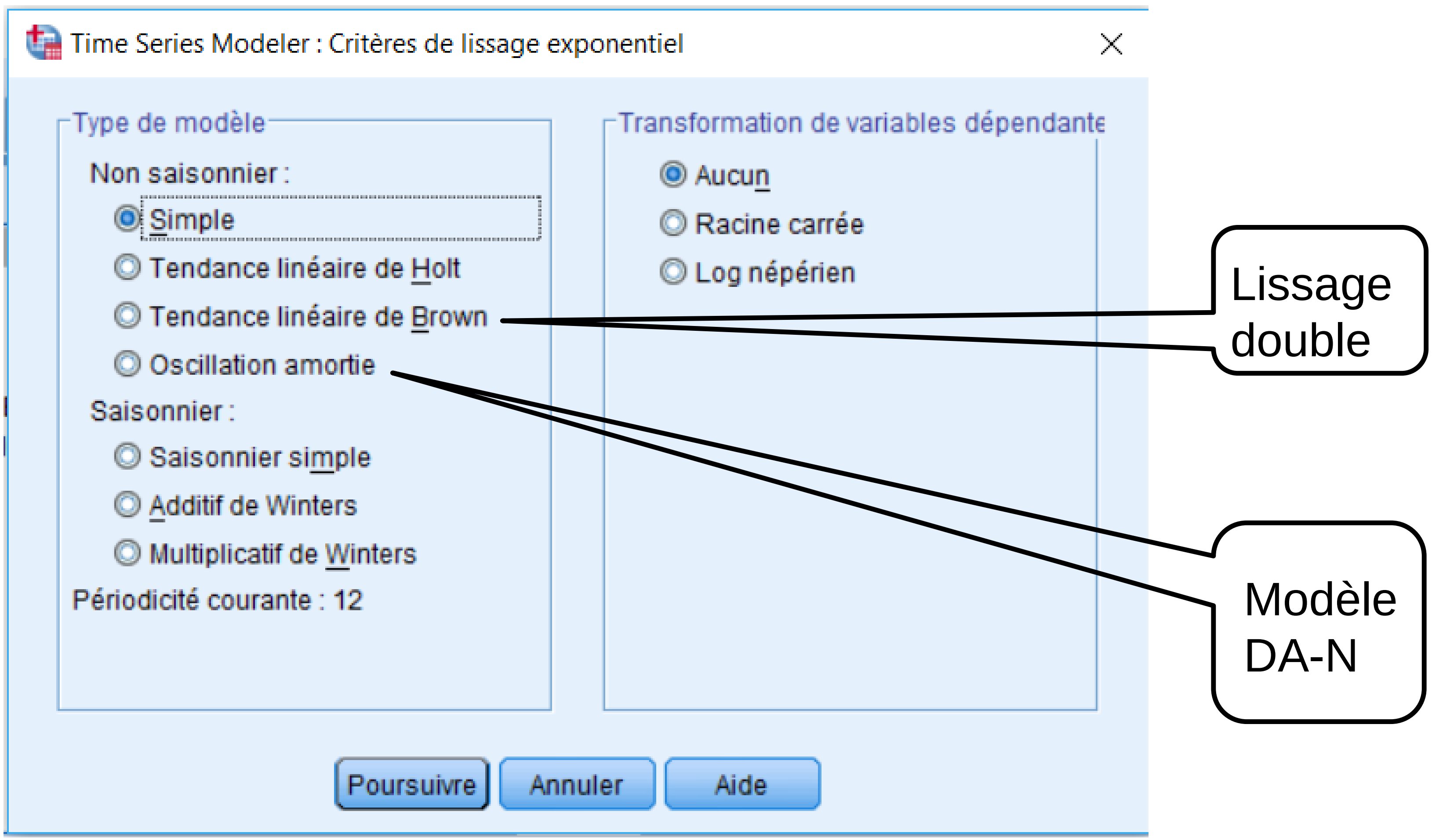
Séries mensuelles ≥ 8*12 = 96

Période d’ajustement : T – 2*s

Période de prévision : 2*s

Le lissage exponentiel avec SPSS

SPSS permet de calculer l'un des modèles suivant



Il permet aussi de choisir le meilleur modèle parmi ces modèles

Les modèles de Hyndman : Approche ETS

E : erreur, T : trend, S : composante saisonnière

Le trend peut prendre les formes suivantes :

None $T_h = l$

Additive $T_h = l + bh$

Additive damped (amortie) $T_h = l + \left(\varphi + \varphi^2 + \cdots + \varphi^h\right)b$

Multiplicative $T_h = lbh$

Multiplicative damped (amortie) $T_h = lb^{\varphi+\varphi^2+\cdots+\varphi^h}$

La composante saisonnière peut prendre les 3 formes suivantes :

N : absence de composantes saisonnières

A : composante saisonnière additive

M : composante saisonnière multiplicative

Trend	Composante saisonnière		
	N (None	A (Additive)	M (Multiplicative)
N None	N,N	N,A	N,M
A Additive	A,N	A,A	A,M
A_d Additive amortie	A_d,N	A_d,A	A_d,M
M Multiplicative	M,N	M,A	M,M
M_d Multiplicative amortie	M_d,N	M_d,A	M_d,M

Il y a 15 possibilités pour combiner le Trend avec la Composante saisonnière, voir fichier ‘‘file’’

L’erreur E peut être combinée soit d’une manière additive ou multiplicative

Hyndman a développé ainsi 30 modèles possibles

Un modèle à erreur additive, a la même formule de prévision que le même modèle à erreur multiplicative

Exemples :
A,N,N : Lissage exponentiel simple
A,A,N : Lissage de Holt
M,A,M : Lissage de H-W multiplicatif avec erreur multiplicative

Le modèle espace d’état linéaire

$$\begin{cases} y_t = w' \ x_{t-1} + \varepsilon_t & \text{équation de mesure} \\ x_t = F x_{t-1} + g\varepsilon_t & \text{équation d'état ou équation de transition} \end{cases}$$

y_t une série chronologique, ε_t est bruit blanc

x_t vecteur de taille m, appelé vecteur d’état

F matrice carrée de taille (m,m), g est un vecteur de taille (m,1)

w’ est un vecteur de taille (1,m)

Le modèle espace d'état non linéaire

$$\begin{cases} y_t = w(x_{t-1}) + r(x_{t-1})\varepsilon_t \\ x_t = f(x_{t-1}) + g(x_{t-1})\varepsilon_t \end{cases}$$

y_t une série chronologique, ε_t est bruit blanc

x_t vecteur de taille m, appelé vecteur d'état

w et r sont deux fonctions de $R^m \to R$

f et g sont deux fonctions de $R^m \to R^m$

Le fichier ''file modèle espace d'état'' reprend les 30 équations différentes des modèles ETS

On pose $\mu_t = w(x_{t-1})$

Si le modèle est à erreur additive alors r$(x_{t-1}) = 1 \to y_t = \mu_t + \varepsilon_t$

Si le modèle est à erreur multiplicative alors r$(x_{t-1}) = \mu_t$

$$\to dans\ ce\ cas\ y_t = \mu_t(1 + \varepsilon_t)$$

Calcul de la fonction vraisemblance du modèle espace d'état linéaire

$$\begin{cases} y_t = w' \, x_{t-1} + \varepsilon_t \\ x_t = F x_{t-1} + g\varepsilon_t \end{cases}$$

On suppose que le bruit ε_t est normale
x_0 est le vecteur d'état initial

x_t est une combinaison de x_0 et $\varepsilon_1, \dots, \varepsilon_t$, il en est de même pour y_t

$$E(y_t/x_{t-1}) = w'x_{t-1} \text{ et } V(y_t/x_{t-1}) = V(\varepsilon_t) = \sigma^2$$

La fonction vraisemblance est la densité de probabilité du vecteur aléatoire $(y_1, \dots, y_n)$: $p(y/x_0)$

$$p(y/x_0) = \prod_{t=1}^{n} p(y_t/y_{t-1}, \dots, y_1, x_0)$$

La densité conjointe s'écrit comme produit de densités conditionnelles

n étant le nombre de valeurs disponibles de la série temporelle

$$p(y/x_0) = \prod_{t=1}^{n} p(y_t/y_{t-1}, \dots, y_1, x_0) = \prod_{t=1}^{n} p(y_t/x_{t-1})$$

La composante aléatoire de y_t si on fixe $y_{t-1}, \dots, y_1, x_0$ est la même que celle de y_t si on fixe x_{t-1}

Si $\varepsilon_t \to N(0, \sigma)$ alors $y_t/x_{t-1} \to N(w' x_{t-1}, \sigma)$

$$f_{y_t/x_{t-1}}(y_t) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(y_t - w' x_{t-1})^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{\varepsilon_t^2}{2\sigma^2}\right)$$

$$p(y/x_0) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\varepsilon_t^2}{2\sigma^2}} = (2\pi\sigma^2)^{-n/2} exp\left(\sum_{t=1}^{n} -\frac{\varepsilon_t^2}{2\sigma^2}\right)$$

Calcul de la fonction vraisemblance du modèle espace d'état non linéaire

$$\begin{cases} y_t = w(x_{t-1}) + r(x_{t-1})\varepsilon_t \\ x_t = f(x_{t-1}) + g(x_{t-1})\varepsilon_t \end{cases}$$

On suppose que le bruit ε_t est normale
x_0 est le vecteur d'état initial

$$p(y/x_0) = \prod_{t=1}^{n} p(y_t/y_{t-1}, \dots, y_1, x_0) = \prod_{t=1}^{n} p(y_t/x_{t-1})$$

Si $\varepsilon_t \to N(0, \sigma)$ alors $y_t/x_{t-1} \to N(w(x_{t-1}), |r(x_{t-1})|\sigma)$

$$p(y/x_0) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi}|r(x_{t-1})|\sigma} e^{-\frac{\varepsilon_t^2}{2\sigma^2}}$$

$$= (2\pi\sigma^2)^{-n/2} \prod_{t=1}^{n} \frac{1}{|r(x_{t-1})|} exp\left(\sum_{t=1}^{n} -\frac{\varepsilon_t^2}{2\sigma^2}\right)$$

On a utilisé le fait que : $\varepsilon_t = \frac{y_t - w(x_{t-1})}{r(x_{t-1})}$

Le logarithme de la fonction vraisemblance : $Ln(\theta, x_0, \sigma^2)$

θ est le vecteur des paramètres du modèle espace d'état

$$LnL(\theta, x_0, \sigma^2) = -\frac{n}{2}Ln(2\pi\sigma^2) - \sum_{t=1}^{n} Ln(|r(x_{t-1})|) - \frac{1}{2}\sum_{t=1}^{n}\frac{\varepsilon_t^2}{\sigma^2}$$

$$\frac{\partial LnL(\theta, x_0, \sigma^2)}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{t=1}^{n}\frac{\varepsilon_t^2}{\sigma^3} = 0 \rightarrow \hat{\sigma}^2 = \sum_{t=1}^{n}\frac{\varepsilon_t^2}{n}$$

$$LnL(\theta, x_0, \hat{\sigma}^2) = -\frac{n}{2}Ln(2\pi\hat{\sigma}^2) - \sum_{t=1}^{n} Ln(|r(x_{t-1})|) - \frac{1}{2}\sum_{t=1}^{n}\frac{\varepsilon_t^2}{\hat{\sigma}^2}$$

$$-2LnL(\theta, x_0, \hat{\sigma}^2) = nLn(2\pi\hat{\sigma}^2) + 2\sum_{t=1}^{n} Ln(|r(x_{t-1})|) + n$$

$$-2LnL(\theta, x_0, \hat{\sigma}^2) = nLn(2\pi e\hat{\sigma}^2) + 2\sum_{t=1}^{n} Ln(|r(x_{t-1})|)$$

L’estimation des paramètres θ et x_0 se fait en minimisant la grandeur :

$$L^*(\theta, x_0) = nLn\left(\sum_{t=1}^{n} \varepsilon_{\mathrm{t}}^2\right) + 2\sum_{t=1}^{n} Ln(|r(x_{t-1})|)$$

Cette grandeur devient : $nLn(\sum_{t=1}^{n} \varepsilon_t^2)$ si le modèle espace d’état est linéaire

L’approche ETS consiste à calculer les 30 modèles ETS voir fichier : “file modèle espace d’état”

Le meilleur modèle est celui qui **minimise** un critère de sélection

Forme des critères de sélection IC (Information criterion)

$$IC = L^*(\theta, x_0) + q\zeta(n)$$

$q\zeta(n)$ est la fonction pénalité et q le nombre de paramètres de θ plus le nombre de composantes non stationnaire du vecteur d’état x_t

Exemples de critères de sélection

Critère	$\zeta(n)$
AIC	2
BIC	Ln(n)
AICC	2n/(n-q-1)

Hyndman a séparé les 30 modèles ETS en 5 groupes :

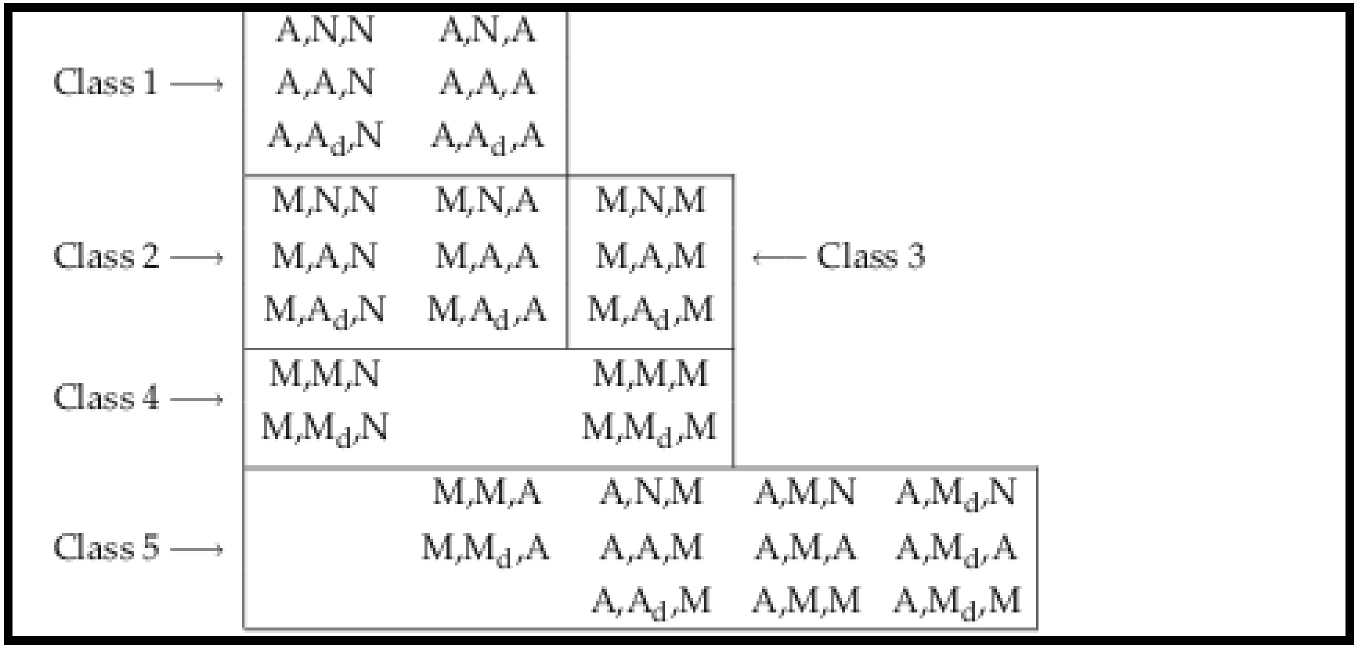
Gr1 : Le modèle espace d'état linéaire et E additive

Gr2 : Le modèle espace d'état linéaire et E multiplicative

Gr3 : MNM, MAM, MA_dM

Gr4 : MMN, MMM, MM_dN, MM_dM

Gr5 : les autres modèles



Classification des modèles ETS en groupes

Calculs des prévisions et intervalles de confiance

Gr1 : Le modèle espace d'état linéaire : E additive et S absente ou additive

$$\begin{cases} y_t = w' \ x_{t-1} + \varepsilon_t \\ x_t = F x_{t-1} + g \varepsilon_t \end{cases}$$

n étant le nombres de valeurs disponibles pour la série temporelle

Pour un horizon h donné, on calcule : $\begin{cases} \hat{y}_{n+h} = E(y_{n+h}/x_n) \\ V(y_{n+h}/x_n) \end{cases}$

Si $\varepsilon_t \to N(0, \sigma^2)$, alors $y_{n+h}/x_n \to N\left(w' F^{h-1} x_n, B\sigma^2\right)$

En effet : $E(x_{n+1}/x_n) = E(F x_n + g \varepsilon_{n+1}/x_n) = F x_n$

Puisque x_n est une combinaison de x_0 et $\varepsilon_1, \dots, \varepsilon_n$, alors ε_{n+1} et x_n sont indépendants : $E(\varepsilon_{n+1}/x_n) = E(\varepsilon_{n+1}) = 0$

De même on démontre que $E(x_{n+h}/x_n) = F^h x_n$

Calcul de $V(x_{n+h}/x_n)$

$V(x_{n+1}/x_n) = V(Fx_n + g\varepsilon_{n+1}/x_n) = V(g\varepsilon_{n+1}/x_n)$ car $V(Fx_n/x_n) = V(cst) = 0$

et ε_{n+1} et x_n sont indépendants

Donc $V(g\varepsilon_{n+1}/x_n) = V(g\varepsilon_{n+1}) = gg'V(\varepsilon_{n+1}) = gg'\sigma^2$, gg' étant une

matrice carré d'ordre m

$$V(x_{n+2}/x_n) = V(Fx_{n+1} + g\varepsilon_{n+2}/x_n) = V(Fx_{n+1}/x_n) + V(g\varepsilon_{n+2}/x_n)$$
$$= FV(x_{n+1}/x_n)F' + gg'\sigma^2 = (Fgg'F' + gg')\sigma^2$$

$$V(x_{n+h}/x_n) = V(Fx_{n+h-1} + g\varepsilon_{n+h}/x_n) = V(Fx_{n+h-1}/x_n) + V(g\varepsilon_{n+h}/x_n)$$
$$= \left(\sum_{i=0}^{h-1} F^i gg'F^{i'}\right)\sigma^2 \text{ si } h \geq 2$$

$$V(y_{n+h}/x_n) = V(w'x_{n+h-1} + \varepsilon_{n+h}/x_n) = w'V(x_{n+h-1}/x_n)w + V(\varepsilon_{n+h}/x_n)$$
$$= \left(w'\left(\sum_{i=0}^{h-2} F^i gg'\, F^{i'}\right)w + 1\right)\sigma^2 \text{ si } h \geq 2$$

$$V(y_{n+1}/x_n) = V(w'x_n + \varepsilon_{n+1}/x_n) = V(\varepsilon_{n+1}/x_n) = \sigma^2$$

Pour un horizon h donné, la loi de y_{n+h}/x_n suit une loi normale $N\left(w'F^{h-1}x_n, B\sigma^2\right)$

$$B = \begin{cases} 1 \quad si\ h = 1 \\ w'F^i gg'\ F^{i'} w + 1 \qquad si\ h \geq 2 \end{cases}$$

Connaissant la loi de y_{n+h}/x_n, on détermine un intervalle de confiance de niveau α : w$'F^{h-1}x_n \pm t_\alpha\sqrt{B}\sigma$ avec $P(U > t_\alpha) = \frac{\alpha}{2}$ et $U \rightarrow N(0,1)$

Gr2 : Le modèle espace d'état linéaire avec E multiplicative et S absente ou additive

Le modèle espace d'état de ce groupe a la forme suivante :

$$\begin{cases} y_t = w'x_{t-1}(1 + \varepsilon_t) \\ x_t = (F + gw'\varepsilon_t)x_{t-1} \end{cases}$$

On a pour un horizon h, $\hat{y}_{n+h} = E(y_{n+h}/x_n)$

$E(y_{n+1}/x_n) = \mathrm{E}(w'x_n(1+\varepsilon_{n+1})/x_n) = w'x_n$, car x_n et ε_{n+1} sont indépendantes

$$E(y_{n+h}/x_n) = \mathrm{E}((w'x_{n+h-1} + w'x_{n+h-1}\varepsilon_{n+h})/x_n) = E(w'x_{n+h-1}/x_n)$$

$$E(x_{n+h}/x_n) = E(Fx_{n+h-1} + gw'x_{n+h-1}\varepsilon_{n+h}/x_n) = FE(x_{n+h-1}/x_n) = F^h x_n$$

$$E(y_{n+h}/x_n) = \mathrm{w}'\mathrm{E}(x_{n+h-1}/x_n) = w'F^{h-1}x_n$$

<u>Calcul de $V(y_{n+h}/x_n)$</u>

$$V(y_{n+1}/x_n) = V(w'x_n(1+\varepsilon_{n+1})/x_n) = w'x_n x_n'\, w\sigma^2$$

Si on suppose que ε_i, i = 1,…,n → $N(0,\sigma)$ alors $y_{n+1}/x_n \rightarrow N(w'x_n, w'x_n x_n'\, w\sigma^2)$

$$y_{n+h}/x_n = w'x_{n+h-1}(1+\varepsilon_{n+h})/x_n$$

On remarque que y_{n+h}/x_n n'est pas une loi normale car c'est le produit de 2 fonctions de lois normales indépendantes

$$V(y_{n+h}/x_n) = V(w'x_{n+h-1}/x_n) + V(w'x_{n+h-1}\varepsilon_{n+h}/x_n) +$$
$$2COV(w'x_{n+h-1}, w'x_{n+h-1}\varepsilon_{n+h}/x_n)$$

$$COV(w'x_{n+h-1}, w'x_{n+h-1}\varepsilon_{n+h}/x_n) = w'COV(x_{n+h-1}, x_{n+h-1}\varepsilon_{n+h}/x_n)w$$

$$COV(x_{n+h-1}, x_{n+h-1}\varepsilon_{n+h}/x_n) = E(x_{n+h-1}x'_{n+h-1}\varepsilon_{n+h}/x_n) -$$
$$E(x_{n+h-1}/x_n)E(x_{n+h-1}\varepsilon_{n+h}/x_n)' = \mathrm{E}(x_{n+h-1}x'_{n+h-1}/x_n)E(\varepsilon_{n+h}/x_n) = 0$$

$$V(w'x_{n+h-1}\varepsilon_{n+h}/x_n) = w'V(x_{n+h-1}\varepsilon_{n+h}/x_n)\mathrm{w}$$

$$V(x_{n+h-1}\varepsilon_{n+h}/x_n)$$
$$= E(x_{n+h-1}\varepsilon_{n+h}\varepsilon_{n+h}x'_{n+h}/x_n) - E(x_{n+h-1}\varepsilon_{n+h}/x_n)E(x_{n+h-1}\varepsilon_{n+h}/x_n)'$$
$$= \mathrm{E}(x_{n+h-1}x'_{n+h-1}/x_n)E\left(\varepsilon^2_{n+h}/x_n\right)$$
$$= (V(x_{n+h-1}/x_n) + E(x_{n+h-1}/x_n)E(x_{n+h-1}/x_n)')\sigma^2$$
$$= \left(V(x_{n+h-1}/x_n) + F^{n+h-1}x_nx'_nF^{n+h-1'}\right)\sigma^2$$

Donc $V(y_{n+h}/x_n) =$

$$w'V(x_{n+h-1}/x_n)w + w'\left(V(x_{n+h-1}/x_n) + F^{n+h-1}x_n x_n' F^{n+h-1'}\right)w\sigma^2$$

$$V(y_{n+h}/x_n) = w'V(x_{n+h-1}/x_n)w(1+\sigma^2) + w'F^{n+h-1}x_n x_n' F^{n+h-1'}w\sigma^2$$

Calculons $V(x_{n+h}/x_n) = V_{n+h}$

On a $x_{n+h} = Fx_{n+h-1} + gw'x_{n+h-1}\varepsilon_{n+h}$

$$V_{n+h} = FV_{n+h-1}F' + V(gw'x_{n+h-1}\varepsilon_{n+h}/x_n) + 2COV(Fx_{n+h-1}, gw'x_{n+h-1}\varepsilon_{n+h}/x_n)$$

Comme précédemment, la covariance est nulle

$$V_{n+h} = FV_{n+h-1}F' + gw'V(x_{n+h-1}\varepsilon_{n+h}/x_n)\mathrm{wg}'$$

$$= FV_{n+h-1}F' + gw'\left(V_{n+h-1} + F^{n+h-1}x_n x_n' F^{n+h-1'}\right)wg'\sigma^2$$

$$V_{n+1} = V(x_{n+1}/x_n) = V(Fx_n + gw'x_n\varepsilon_{n+1}/x_n) = gw'x_n x_n' wg'\sigma^2$$

Algorithme de calcul de $V(y_{n+h}/x_n)$ pour $h \geq 2$

$$V_{n+1} = gw'x_nx_n'wg'\sigma^2$$

Pour i ← 2 à h

$$V(y_{n+i}/x_n) = w'V_{n+i-1}w + w'\left(V_{n+i-1} + F^{n+h-1}x_nx_n'F^{n+h-1'}\right)w\sigma^2$$

$$V_{n+i} = FV_{n+i-1}F' + gw'\left(V_{n+i-1} + F^{n+i-1}x_nx_n'F^{n+i-1'}\right)wg'\sigma^2$$

Fin_Pour

Gr3 : Le modèle espace d'état non linéaire avec E multiplicative et S multiplicative

$$\begin{cases} y_t = w_1'x_{t-1}w_2'z_{t-1}(1+\varepsilon_t) \\ x_t = (F_1 + g_1w_1'\varepsilon_t)x_{t-1} \\ z_t = (F_2 + g_2w_2'\varepsilon_t)z_{t-1} \end{cases}$$

x_t : niveau et pente, z_t : composante saisonnière

On a pour un horizon h, $\hat{y}_{n+h} = E(y_{n+h}/x_n\,, z_n) = w_1' x_n w_2' z_n$ si h=1

Il est possible de calculer $V(y_{n+h}/x_n, z_n)$

Pour les G2 et G3, même si y_{n+h}/x_n n'est pas normale, on utilise la formule suivante pour l'intervalle de confiance de niveau de signification α et à horizon h donné :

$$\hat{y}_{n+h} \pm t_\alpha \sqrt{V(y_{n+h}/x_n)} \text{ avec } P(U > t_\alpha) = \frac{\alpha}{2} \text{ et } U \to N(0,1)$$

Pour G3 on a : $\hat{y}_{n+h} \pm t_\alpha \sqrt{V(y_{n+h}/x_n, z_n)}$

Groupe 4 et Groupe 5

$\hat{y}_{n+h} = E(y_{n+h}/x_n\,) \to$ voir fichier “file”

$V(y_{n+h}/x_n\,)$ n'a pas pu être calculée

Calcul d'un intervalle de confiance par simulation de niveau de signification α

Cette méthode permet de contourner le calcul de $V(y_{n+h}/x_n)$

On dispose de n valeurs d'une série temporelle $y_1, \dots y_n$, pour laquelle on a calculé un modèle ETS

$$\begin{cases} y_t = w(x_{t-1}) + r(x_{t-1})\varepsilon_t \\ x_t = f(x_{t-1}) + g(x_{t-1})\varepsilon_t \end{cases}$$

En estimant les paramètres du modèle et en en considérant le fait que pour t = 1,…,n, $\varepsilon_t = \frac{y_t - w(x_{t-1})}{r(x_{t-1})}$, on calcule x_n

Pour un horizon h donné, on génère h valeurs de la loi $N(0, \hat{\sigma}^2)$, ce qui permet de calculer $y_{n+1}, \dots y_{n+h}$

Cette opération est répétée T fois, T = 5000, 10000, 20000

On obtient ainsi un échantillon de la variable y_{n+h}

Cette échantillon est trié selon l'ordre croissant

LimInf = T$\alpha/2$ et LimSup = T$(1-\alpha/2)$ ième valeurs de l'échantillon trié

Les critères de comparaison

$$MSE = \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_{i/i-1}\right)^2}{n} \; avec \; \hat{y}_{i/i-1} = E(y_i/x_{i-1}) = w(x_{i-1})$$

$$AME = \frac{\sum_{i=1}^{n} abs\left(y_i - \hat{y}_{i/i-1}\right)}{n}$$

On définit $MSE_k = \frac{\sum_{i=k}^{n}\left(y_i - \hat{y}_{i/i-k}\right)^2}{n}$, k = 1,2,3

On a $AMSE = \frac{MSE_1 + MSE_2 + MSE_3}{3}$

La fonction ets du package forecast

Elle permet de calculer le modèle ets d'une série temporelle

Par défaut les paramètres sont estimés par maximisation de la fonction vraisemblance

Il est possible d'estimer les paramètres du modèle en minimisant :

$$\begin{cases} mse \\ amse \\ mae \\ \sigma^2 \end{cases}$$

Le meilleur modèle choisi est celui qui minimise un critère de sélection :

$$\begin{cases} AIC \\ BIC \\ AICC \end{cases}$$

Codes possibles

ets(z) est équivalent à ets(z, model =“ZZZ”)

ets(z, modele = “AMM”) → calculer le model AMM

ets(z, model = “AMM”, damped = TRUE) → calculer le model AM_dM

ets(z, model = “ZMA”, damped = TRUE) → calculer le meilleur model parmi les modèle T = M_d et S = A

A défaut, la tendance est additive

ets(z, allow.multiplicative.trend = TRUE)

ets(z, opt.crit = “amse”, ic = “bic”), calculer le meilleur modèle en minimisant le critère amse pour estimer les paramètres et en utilisant bic pour choisir le meilleur

ets(z, model = “AAA”, opt.crit = “sigma”) calculer le modèle AAA en estimant les paramètres en minimisant σ l’écart type du bruit ε_t