Lissages Exponentiels

MAP-STA2 : Séries chronologiques

Yannig Goude yannig.goude@edf.fr

2018-2019

Les méthodes de lissages exponentielles, dont il existe plusieurs variantes que nous présentons ici, sont des méthodes empiriques de prévision de série temporelle. Elles présentent l'intérêt d'être facilement compréhensibles et leur implémentation récursive en font un outil efficace pour le traitement de gros volumes de données ou dans des systêmes embarqués disposant de peu de mémoire. Il faut noter que, bien que largement utilisés, ces méthodes souffrent d'un manque de bases théoriques solides comme celles des ARMA, ARIMA et SARIMA.

toutes ces méthodes consistent à ajuster à une chronique de série temporelle une estimation locale de ce que va être sa valeur future. Selon les variantes:

- une constante pour le lissage exponentiel simple
- une droite pour le lissage exponentiel double ou de Holt
- des fonctions polynomiales ou périodiques pour les lissages plus généraux

Lissage exponentiel simple

Un algorithme de base pour la prévision de séries temporelles univariées est le lissage exponentiel, c'est la plus ancienne des méthodes que nous verrons dans ce chapitre.

On peut voir le lissage exponentiel comme une méthode de prévision mais également, comme son nom l'indique, comme une technique de lissage de données.

définition soit une série temporelle y_t. On appelle lissage exponentiel simple de paramètre $\alpha \in [0,1]$ de cette série le processus $\widehat{y}_t$ définie ainsi:

$$\widehat{y}_{t+1/t} = \alpha y_t + (1-\alpha)\widehat{y}_{t/t-1}$$

on a donc:

$$\widehat{y}_{t+1/t} = \sum_{i=0}^{t-1} \alpha(1-\alpha)^i y_{t-i}$$

la prévision de l'instant $t+1$ est donc une somme pondérée des valeurs passées de la série, les poids décroissant exponentiellement dans le passé. La *mémoire* de la prévision dépend de α. Plus α est proche de 1 plus les observations récentes influent sur la prévision, à l'inverse un α proche de 0 conduit à une prévision très stable prenant en compte un passé lointain.

Une autre façon d'écrire le lissage exponentiel (error correction form):

$$\widehat{y}_{t+1/t} = \widehat{y}_{t/t-1} + \alpha(y_t - \widehat{y}_{t/t-1})$$

remarque $\widehat{y}_{t+1/t}$ est une prévision à horizon 1. Il est parfois nécessaire d'effectuer une prévision à un horizon h quelconque. On notera par la suite $\widehat{y}_{t+h/t}$ la prévision de y_{t+h} conditionellement à $(y_1, ..., y_t)$. Pour le lissage exponentiel simple cette prévision est tout simplement: $\widehat{y}_{t+h/t} = \widehat{y}_{t+1}$ car on approxime le futur de la série à une constante (cf remarque).

remarque le prédicteur obtenu par lissage exponentiel simple peut être vu comme celui qui minimise un problème de moindres carrés pondérés:

En effet, si on note:

$$\widehat{a} = \text{argmin}_a \sum_{i=0}^{t-1} (1-\alpha)^i (y_{t-i} - a)^2$$

on a $\widehat{a} = \sum_{i=0}^{t-1}(1-\alpha)^i y_{t-i} / \sum_{i=0}^{t-1}(1-\alpha)^i$, soit

$\widehat{a} = \frac{\alpha}{1-(1-\alpha)^t} \sum_{i=0}^{t-1}(1-\alpha)^i y_{t-i}$

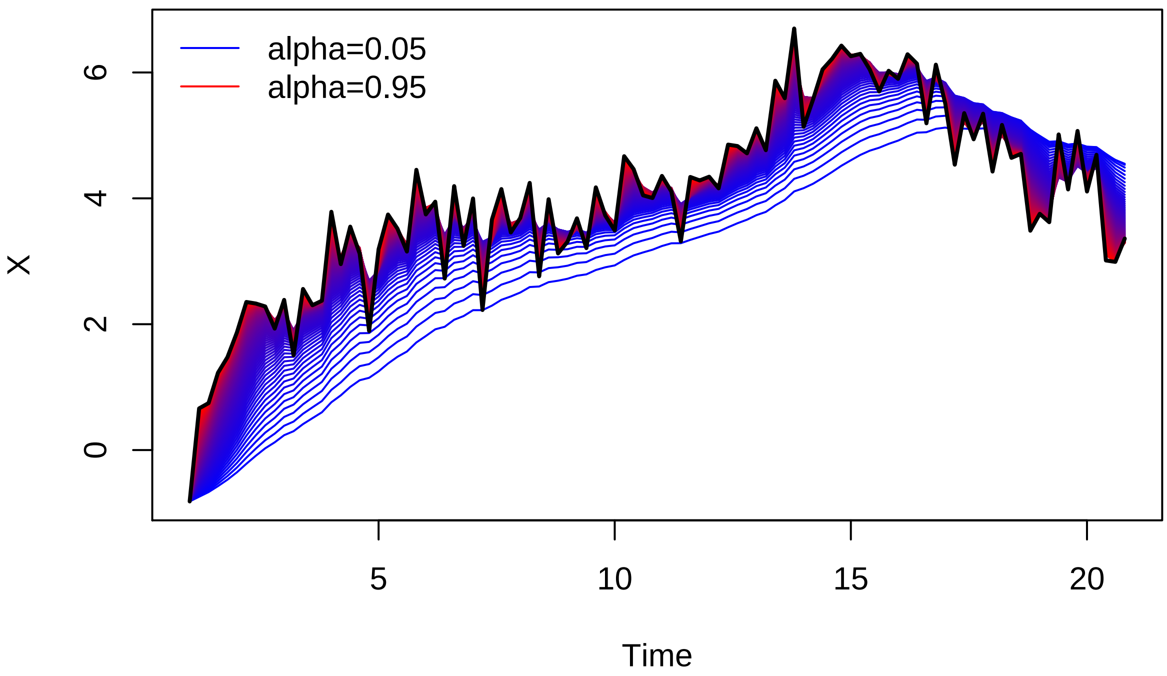
ce qui, si t est suffisamment grand tend vers l'estimateur de lissage exponentiel simple définie précédemment.

Si on défnie y_t comme $(0, ..., 0, y_1, ..., y_t)$ ie valant 0 pour les indices temporelles inférieurs à 1 alors

$$\widehat{y}_{t+1} = \text{argmin}_a \sum_{i=0}^{\infty} (1-\alpha)^i (y_{t-i} - a)^2$$

Voilà un exemple de lissage exponentiel simple sur des données simulées de la façon suivante:

```
set.seed(150)
n<-100
t<-c(1:n)
eps<-rnorm(n,0,1/2)
T<-log(t)+pmax(t-50,0)/10-pmax(t-70,0)/5
X<-T+eps
```



remarque en pratique il est bien sur nécessaire d'initialiser le lissage. Une possibilité est d'attribuer $\widehat{y}_{1/0} = y_1$. On peut également optimiser cette constante d'initialisation sur les données.

Lissage exponentiel double (ou de Holt)

Holt (1957) a étendu le lissage exponentiel simple au cas du lissage exponentielle linéaire. L'idée est d'ajuster une droite au lieu d'une constante dans l'approximation locale de la série.

définition soit une série temporelle y_t. On appelle lissage exponentiel double (ou de Holt) de paramètre $\alpha \in [0,1]$ de cette série le processus $\widehat{y}_t$ définie ainsi:

$$\widehat{y}_{t+h/t} = l_t + b_t h$$

avec:

$$\begin{cases} l_t = l_{t-1} + b_{t-1} + (1-(1-\alpha)^2)(y_t - \widehat{y}_{t/t-1}) \\ b_t = b_{t-1} + \alpha^2 (y_t - \widehat{y}_{t/t-1}) \end{cases}$$

l_t et b_t minimise à chaque instant:

$$(\widehat{l}, \widehat{b}) = \text{argmin}_{l,b} \sum_{i=0}^{\infty} (1-\alpha)^i (y_{t-i} - (l+bi))^2$$

preuve

on note

$$C(b,l) = \sum_{i=0}^{\infty} (1-\alpha)^i (y_{t-i} - (l-bi))^2$$

en dérivant C en l et b on obtient:

$$\frac{\partial C}{\partial l} = -2 \sum_{i=0}^{\infty} (1-\alpha)^i (y_{t-i} - l + bi)$$

$$\frac{\partial C}{\partial b} = 2 \sum_{i=0}^{\infty} i(1-\alpha)^i (y_{t-i} - l + bi)$$

le minimum se réalisant ou les dérivées s'annulent (fonction convexe) et en remarquant que $\sum_{i=0}^{\infty}(1-\alpha)^i = 1/\alpha$, $\sum_{i=0}^{\infty} i(1-\alpha)^i = \frac{1-\alpha}{\alpha^2}$, $\sum_{i=0}^{\infty} i^2(1-\alpha)^i = \frac{(1-\alpha)(2-\alpha)}{\alpha^3}$, on a:

$$\begin{cases} \sum_{i=0}^{\infty}(1-\alpha)^i y_{t-i} - \frac{l}{\alpha} + \frac{b(1-\alpha)}{\alpha^2} = 0 \\ \sum_{i=0}^{\infty} i(1-\alpha)^i y_{t-i} - l\frac{1-\alpha}{\alpha^2} + b\frac{(1-\alpha)(2-\alpha)}{\alpha^3} = 0 \end{cases}$$

soit

$$\begin{cases} \alpha \sum_{i=0}^{\infty}(1-\alpha)^i y_{t-i} - l + \frac{b(1-\alpha)}{\alpha} = 0 \\ \alpha^2 \sum_{i=0}^{\infty} i(1-\alpha)^i y_{t-i} - l(1-\alpha) + b\frac{(1-\alpha)(2-\alpha)}{\alpha} = 0 \end{cases}$$

on note: $L_1(t) = \alpha \sum_{i=0}^{\infty}(1-\alpha)^i y_{t-i}$ et $L_2(t) = \alpha \sum_{i=0}^{\infty}(1-\alpha)^i L_1(t-i)$

remarquons que $L_2(t) - \alpha L_1(t) = \alpha^2 \sum_{i=1}^{\infty} i(1-\alpha)^i y_{t-i}$ car:

$$L_2(t) - \alpha L_1(t) = \alpha \sum_{i=1}^{\infty}(1-\alpha)^i L_1(t-i)$$

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

$$= \alpha^2 \sum_{i=1}^{\infty}\sum_{k=j}^{\infty}(1-\alpha)^k y_{t-k} = \alpha^2 \sum_{i=1}^{\infty} i(1-\alpha)^i y_{t-i}$$

alors:

$$\begin{cases} L_1(t) - l + \frac{b(1-\alpha)}{\alpha} & =0 \\ L_2(t) - \alpha L_1(t) - l(1-\alpha) + b\frac{(1-\alpha)(2-\alpha)}{\alpha} & =0 \end{cases}$$

et

$$\begin{cases} L_1(t) = & l - \frac{b(1-\alpha)}{\alpha} \\ L_2(t) = & l - 2b\frac{(1-\alpha)}{\alpha} \end{cases}$$

ainsi

$$\begin{cases} l = & 2L_1(t) - L_2(t) \\ b = & \frac{\alpha}{1-\alpha}(L_1(t) - L_2(t)) \end{cases}$$

nous pouvons ensuite en déduire les formules de récurrences.

$$L_1(t) = \alpha y_t + (1-\alpha)L_1(t-1)$$

$$\begin{cases} L_2(t) = & \alpha L_1(t) + (1-\alpha)L_2(t-1) \\ = & \alpha^2 y_t + \alpha(1-\alpha)L_1(t-1) + (1-\alpha)L_2(t-1) \end{cases}$$

ainsi

$$l = (1-(1-\alpha)^2)y_t + (1-\alpha)(2-\alpha)L_1(t-1) - (1-\alpha)L_2(t-1)$$

en incorporant $L_1(t) = l_t - \frac{1-\alpha}{\alpha} b_t$ et $L_2(t) = l_t - b_t\frac{2(1-\alpha)}{\alpha}$, on en déduit après simplification:

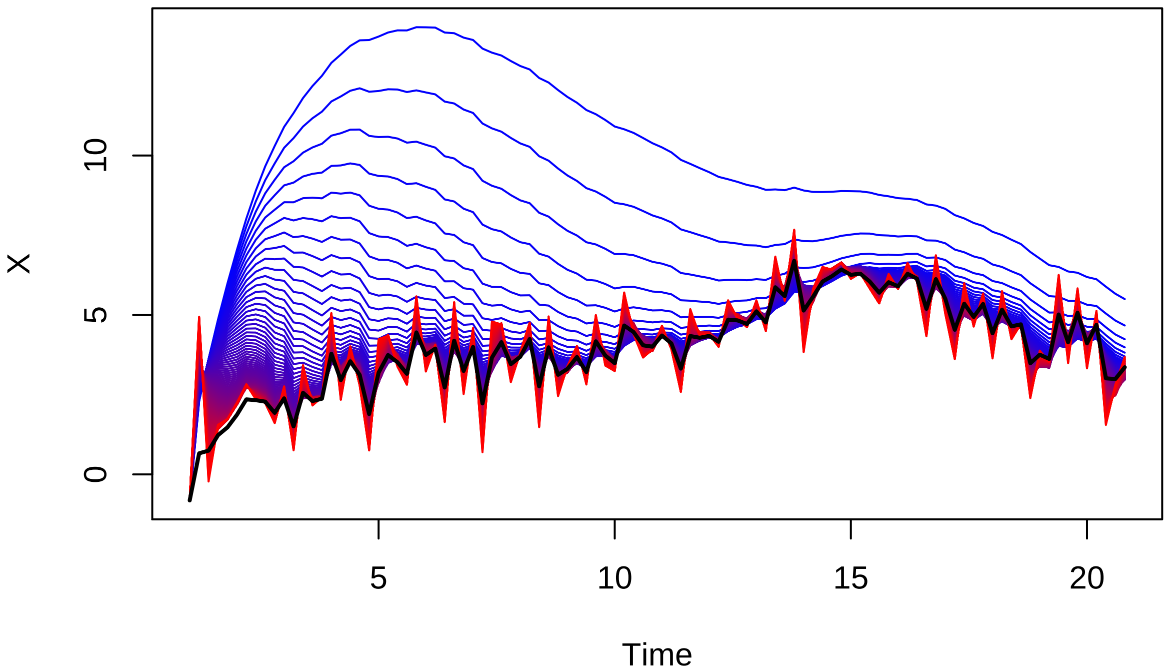
$$l_t = (1-(1-\alpha^2))y_t + (1-\alpha)^2(l_{t-1} + b_{t-1})$$

De même,

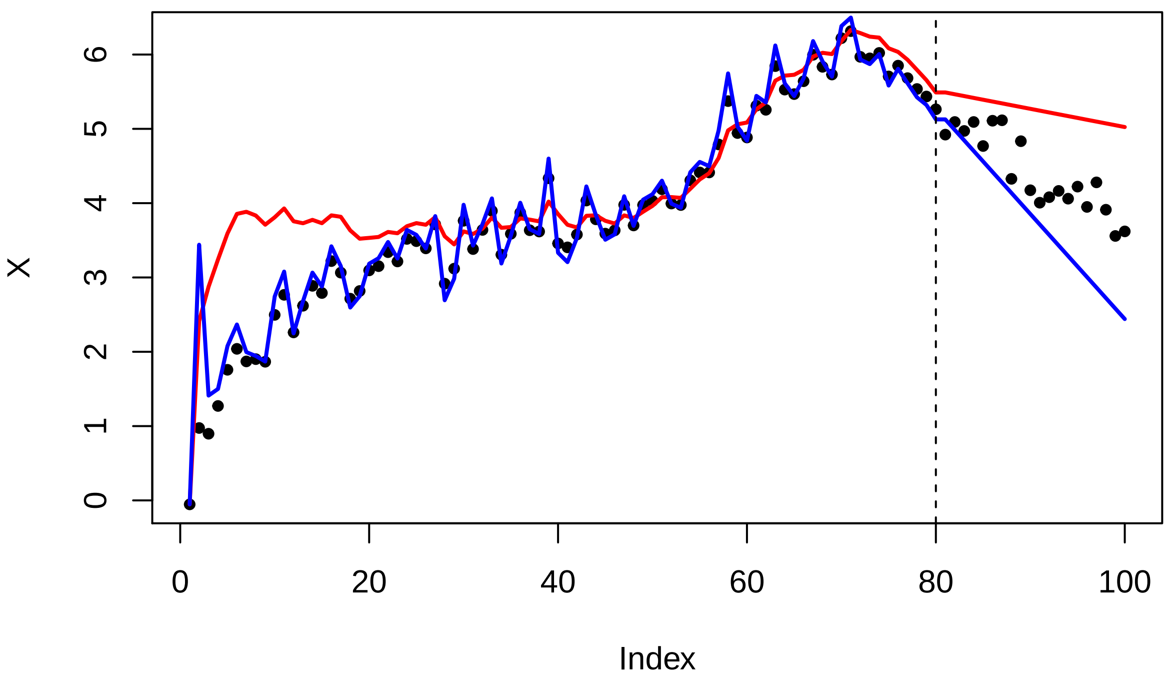
$$b_t = \alpha^2 y_t + \alpha(1-\alpha)L_1(t-1) - \alpha L_2(t-1)$$

$$b_t = \alpha^2 y_t - \alpha^2 l_{t-1} + (1-\alpha^2)b_{t-1}$$

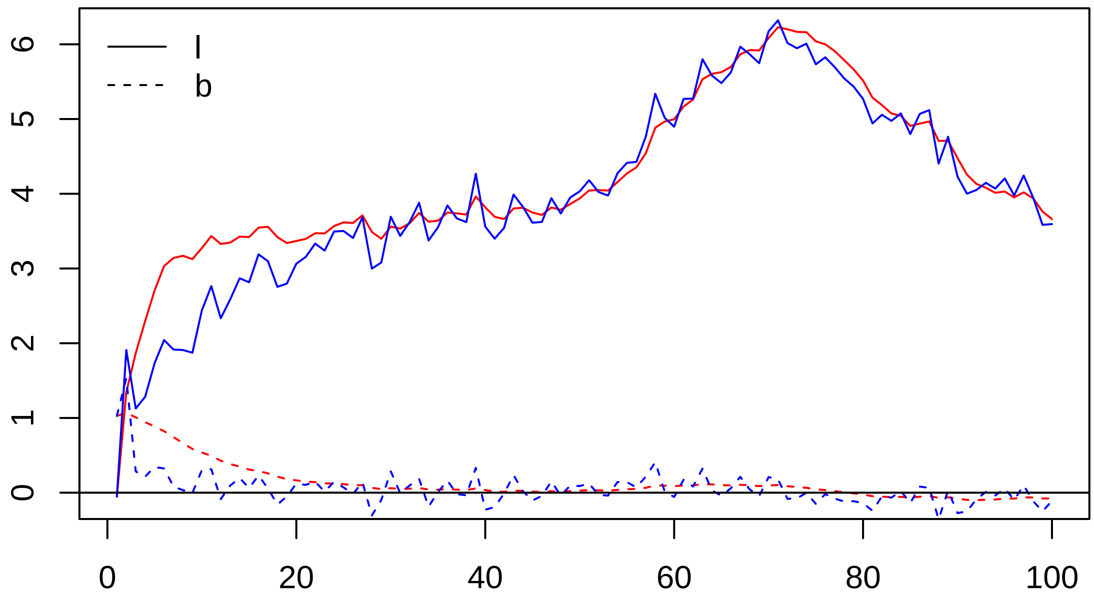
en rappelant que $\widehat{y}_{t+h/t} = l_t + b_t h$ on retrouve les formules de mise à jour attendues.



un exemple de prévisions à horizon $1, 2, ..., 20$:



ainsi que les paramètres d'ordonnées à l'origine et de pente associés:



Lissage exponentiel de Holt-Winters

Cette approche est une généralisation du lissage double, qui permet entre autre de proposer les modèles suivants:

- tendance linéaire locale
- tendance linéaire locale + saisonnalité (modèle additif)
- tendance linéaire locale * saisonnalité (modèle multiplicatif)

Dans ce cas 2 paramètres de lissage entrent en jeu et on ajuste au voisinage de t un fonction linéaire $l_t + hb_t$, h étant l'horizon de prévision.

définition soit une série temporelle y_t. On appelle lissage exponentiel double de Holt-Winters de paramètres $\alpha \in [0,1]$ et $\beta \in [0,1]$ de cette série le processus $\widehat{y}_t$ définie ainsi:

$$\widehat{y}_{t+h} = l_t + hb_t$$

avec

$$\left\{ \begin{array}{l} l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}) \\ b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \end{array} \right.$$

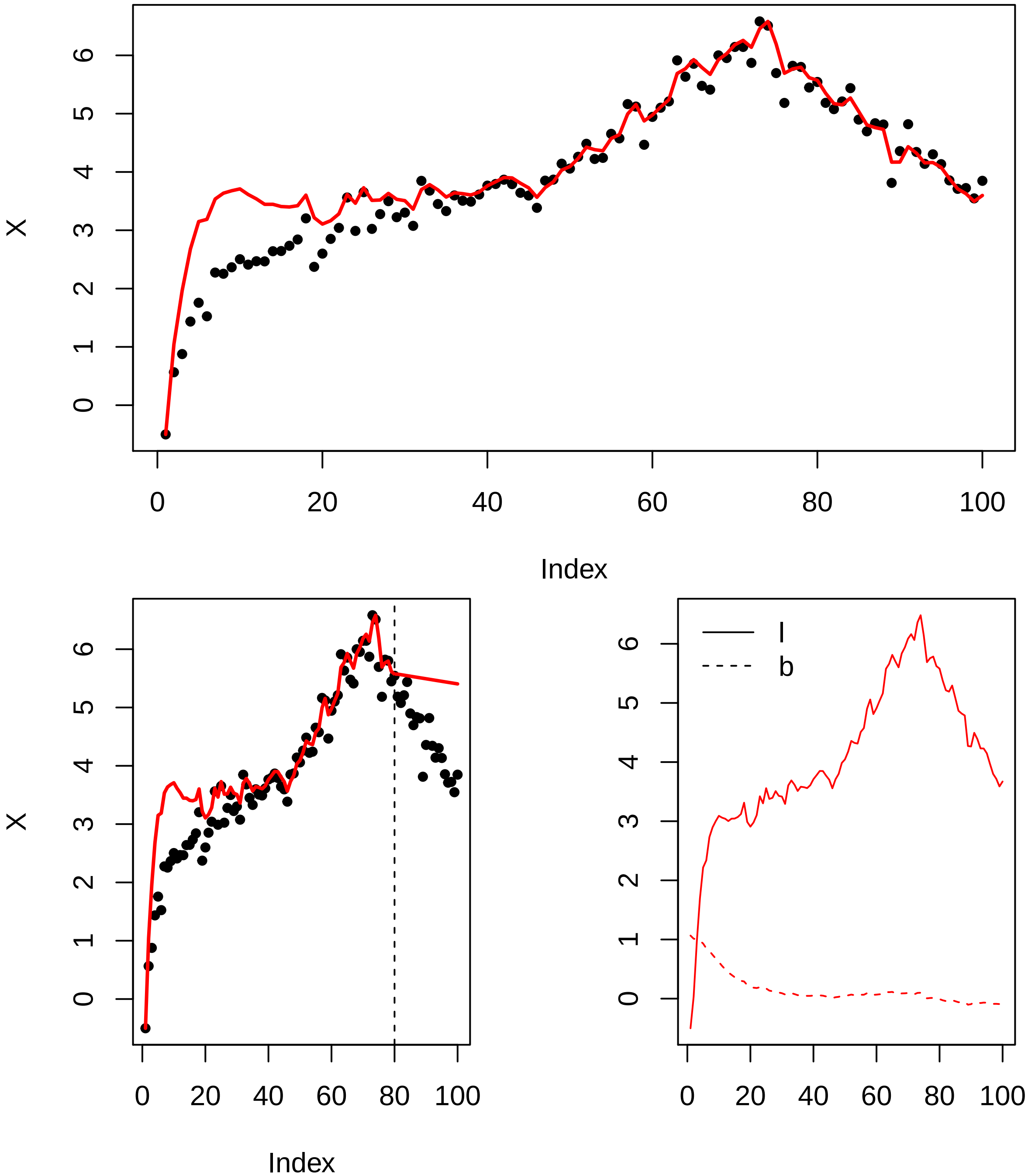
la encore l_t est une estimation du niveau de la série, b_t de sa pente (localement en temps).

remarque la formule de mise à jour du lissage exponentiel double est un cas particulier du lissage de Holt-Winters. En effet on peut écrire dans ce cas:

$$\left\{ \begin{array}{l} l_t = (1-(1-\alpha)^2)y_t + (1-\alpha)^2(l_{t-1} + b_{t-1}) \\ b_t = \frac{\alpha^2}{1-(1-\alpha)^2}(l_t - l_{t-1}) + (1 - \frac{\alpha^2}{1-(1-\alpha)^2})b_{t-1} \end{array} \right.$$

remarque l_t est le barycentre affecté des poids α et $1-\alpha$ de la dernière valeur de y observé et sa prévision à horizon 1. L'algorithme "corrige" donc la prévision de la constante en prenant en compte le dernier écart observé. De même, b_t est au barycentre de la dernière pente prévue et l'écart entre les 2 dernières ordonnées à l'origine prévues.

Sur les données simulées précédamment cela donne, en prenant $\alpha = 0.5$, $\beta = 0.1$:



Pour ajuster une composante saisonnière, on considère le modèle (localement au voisinage de t):

$$\widehat{y}_{t+h} = l_t + hb_t + s_t$$

ou s_t est une composante périodique de période T.

définition soit une série temporelle y_t. On appelle lissage exponentiel double de Holt-Winters saisonnier de paramètres $\alpha \in [0,1]$, $\beta \in [0,1]$, $\delta \in [0,1]$ de cette série le processus $\widehat{y}_t$ définie ainsi:

$$\widehat{y}_{t+h} = l_t + hb_t + s_t$$

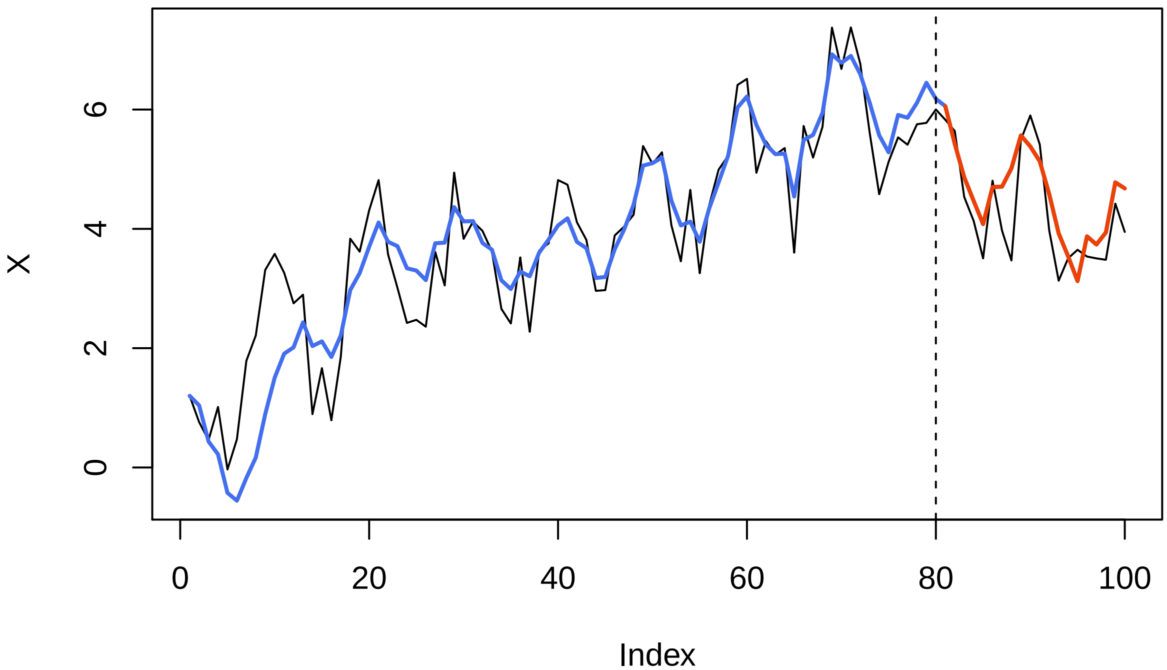
avec

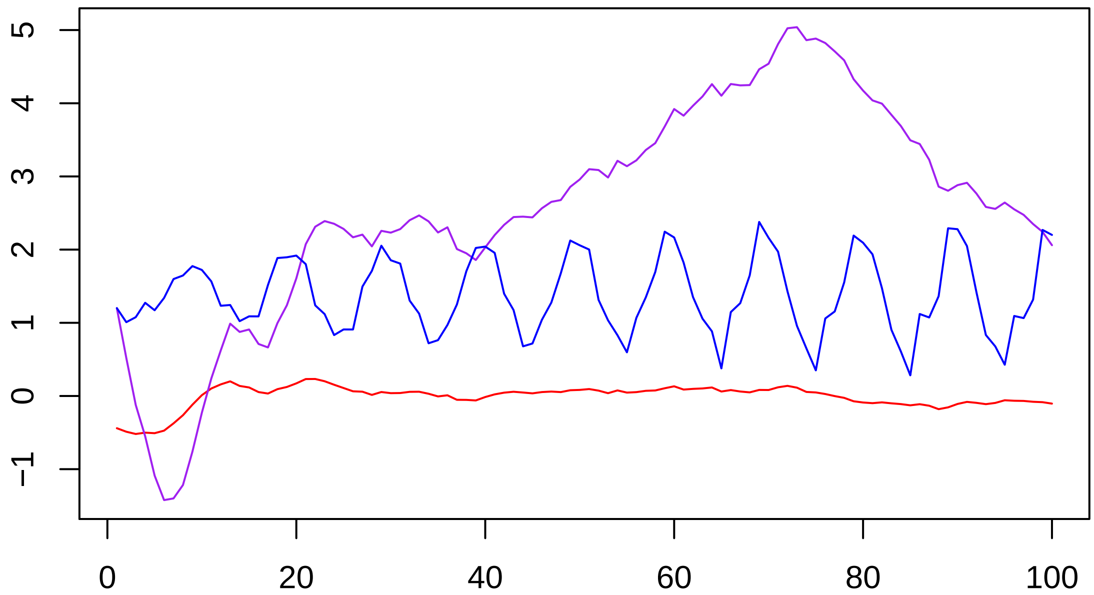
$$\begin{cases} l_t = & \alpha(y_t - s_{t-T}) + (1-\alpha)(l_{t-1} + b_{t-1}) \\ b_t = & \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\ s_t = & \delta(y_t - l_t) + (1-\delta)s_{t-T} \end{cases}$$

Reprenons les données simulées précédamment en ajoutant une composante périodique de période $T = 10$.

```
n<-100
t<-c(1:n)
eps<-rnorm(n,0,1/2)
w<-2*pi/10
S<-cos(w*t)
T<-log(t)+pmax(t-50,0)/10-pmax(t-70,0)/5
X<-T+S+eps
```

Voilà ce qu'on obtient en appliquant un lissage exponentiel saisonnier de paramètres $\alpha = 0.2$, $\beta = 0.2$, $\delta = 0.2$:





Il existe une variante multiplicative de ce lissage saisonnier.

définition soit une série temporelle y_t. On appelle lissage exponentiel double de Holt-Winters saisonnier multiplicatif de paramètres $\alpha \in [0, 1]$, $\beta \in [0, 1]$, $\delta \in [0, 1]$ de cette série le processus $\widehat{y}_t$ définie ainsi:

$$\widehat{y}_{t+h} = (l_t + hb_t)s_t$$

$$\begin{cases} l_t = \alpha \frac{y_t}{s_{t-T}} + (1-\alpha)(l_{t-1} + b_{t-1}) \\ b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\ s_t = \delta \frac{y_t}{l_t} + (1-\delta)s_{t-T} \end{cases}$$

Implémentation en R

Les méthodes de lissage exponentiel sont implémentés en `r` dans la fonction:

```
HoltWinters(x, alpha = NULL, beta = NULL, gamma = NULL,
            seasonal = c("additive", "multiplicative"),
            start.periods = 2, l.start = NULL, b.start = NULL,
            s.start = NULL,
            optim.start = c(alpha = 0.3, beta = 0.1, gamma = 0.1),
            optim.control = list())
```

Il est possible avec cette fonction de fixer les paramètres de lissage, ou (par défaut) de les laisser estimer par la méthode. Dans le cas saisonnier, la période correspond à celle de la série temporelle x (objet `ts`).