

M1 IM, université Nice Sophia Antipolis
Séries temporelles
Sylvain Rubenthaler
<http://math.unice.fr/~rubentha/cours.html>

TP 2 (durée : 3h)

Préliminaires

Créer un fichier texte dans lequel vous répondrez clairement aux questions ci-dessous, en incluant vos codes R, les résultats obtenus sous R (graphique y compris), vos interprétations, remarques ... Une fois ce TP fini, vous mettrez en forme votre compte-rendu et l'exporterez au format pdf (c'est ce qui sera demandé au partiel).

1 Lissage et prévision de données simulées

Simuler les trois séries temporelles suivantes de taille 100 :

- $X_1(t) = \epsilon_t$,
- $X_2(t) = 0.5t + 2\epsilon_t$,
- $X_3(t) = 0.5t + \epsilon_t + 3 \cos(t\pi/6)$,

où les ϵ_t sont i.d.d . de loi $\mathcal{N}(0, 1)$.

1. Pour chaque série, effectuer la prévision $\hat{x}_{t,1}$ ($1 \leq t \leq 99$) par lissage exponentiel simple et double. Tester différentes valeurs du paramètre de lissage α (4 à 5 valeurs), et représenter graphiquement la série ainsi que la prévision. Commenter chaque résultat, et essayer de déterminer graphiquement le lissage le plus adaptée pour chaque série.
2. Calculer pour chaque prévision effectuée la somme des carrés des erreurs, et sélectionner le meilleur modèle à l'aide de cette quantité.
3. Tester maintenant le lissage exponentiel de Holt-Winters avec composante saisonnière additive puis multiplicative (seulement pour les séries avec une composante périodique).
4. Les prédictions obtenues sont-elles meilleures ?

2 Lissage et prévision de la concentration en co2

Le fichier de données co2 contenu dans R¹ contient les concentrations en CO2 à proximité du volcan Mauna Loa (Hawaï) de 1959 à 1997.

Après avoir représenté graphiquement ces données, quel modèle de lissage exponentiel vous semble le mieux approprié ?

Afin de valider ce modèle, tester la prédition des données de 1990 à 1997 en utilisant celles de 1959 à 1989. Si cela vous semble graphiquement correct, utilisez cette méthode pour prédire les concentrations en CO2 de 1997 à 2007. Sinon, tester d'autres méthodes de lissage exponentiel.

3 Lissage et prévision du CAC40

Récupérer le fichier contenant les valeurs de cloture journalière du CAC40 de 1991 à 1998 (données R EuStockMarkets). Essayer de prédire par lissage exponentiel les valeurs de cloture de 1998 en utilisant les valeurs de cloture de 1991 à 1997.

1. Il est dans l'espace de travail dès que vous lancez R.

Appendice : mise en œuvre sous R

La commande `par(mfrow=c(n,p))` permet d'afficher $n*p$ graphique sur la même page.

Obtenir un vecteur aléatoire de longueur n de composantes des variables i.i.d. $\sim \mathcal{N}(0, 1)$: `e=rnorm(n)`. Multiplier/additionner deux vecteurs composante par composante : `x*y`, `x+y`. Prendre la fonction `f` d'un vecteur `x` composante par composante : `f(x)`. Créer un vecteur par concaténation : `v=c()`; `v=c(v,1,2,3)`. Trouver le type d'une variable : `typeof(x)`.

Découper une série temporelle de `c(i,j)` à `c(k,1)` : `xd<-window(x,c(i,j),c(k,1))`.

Les méthodes de lissages exponentiels sont disponibles sous R, grâce à la fonction `HoltWinters`.

Pour une série temporelle `x`, cette procédure permet :

- un lissage exponentiel simple :
`xlisse <- HoltWinters(x, alpha=α, beta=FALSE, gamma=FALSE)`,
- un lissage exponentiel double paramètre α' :
`xlise <- HoltWinters(x, alpha=α, beta=β, gamma=FALSE)` avec $\alpha = 1 - (\alpha')^2$, $\beta = \frac{1-\alpha'}{1+\alpha'}$,
- un lissage de Holt-Winters sans composante saisonnière :
`xlisse <- HoltWinters(x, alpha=α, beta=β, gamma=FALSE)`,
- un lissage Holt-Winters additif :
`xlisse <- HoltWinters(x, alpha=α, beta=β, gamma=γ, seasonal='add')`,
- un lissage Holt-Winters multiplicatif :
`xlisse <- HoltWinters(x, alpha=α, beta=β, gamma=γ, seasonal='mul')`.

À noter que pour un lissage de Holt-Winters avec composante saisonnière la série temporelle `x` doit être un objet de type série temporelle, défini avec la fonction `ts` en précisant la saisonnalité.

Les prévisions à l'horizon `h` sont réalisées à l'aide de la fonction `predict` :

`p<-predict(xlisse, n.ahead=h)` (attention le `n` fait partie de la commande et le `h` est une variable qui peut prendre les valeurs $1, 2, \dots$). Un intervalle de confiance (dont le fondement théorique n'a pas été étudié dans ce cours) peut être obtenu en validant (à `TRUE`) l'option `prediction.interval`. Tracer la prédition et la série sur le même graphique : `xlisse=HoltWinters(x, ...); p=predict(xlisse, n.ahead=50); plot(xlisse, p)`.

Remarque : lorsqu'aucune valeur n'est précisée pour les constantes de lissage, un algorithme interne à la procédure `HoltWinters` se charge d'estimer la meilleure constante possible à partir de la série des observations.