Table des matières

Première partie

Analyse descriptive des séries chronologiques

1 Notations

Définition: Série chronologique

Suite finie de données quantitatives indexée par le temps.
Si on considère une série chronologique de longueur n :
— $t_1,...,t_n$ désigne les n instants successifs d'observation
— y_i sera la valeur mesure à l'instant t_i (en considérant les dates d'observations équidistantes).

2 Modèles de décomposition déterministes

Deux modèles sont étudiés :
1. Le modèle additif
2. Le modèle multiplicatif

combinant chacun :
1. Une tendance f_i
2. Une composante saisonnière s_i
3. Une composante résiduelle e_i

Définition: Modèle additif

Le modèle additif prédit une étiquette sous la forme suivante :

$$y_i = f_i + s_i + e_i,\ i = 1..n$$

avec :

$$\sum_{j=1}^{p} s_j = 0 \text{ et } \sum_{i=1}^{n} e_i = 0$$

Où p désigne une période.
On utilise ce modèle quand, en reliant minima et maxima, on obtient deux droites parallèles.

Définition: Modèle additif

Le modèle multiplicatif prédit une étiquette sous la forme suivante :

$$y_i = f_i(1 + s_i)(1 + e_i),\ i = 1..n$$

avec :

$$\sum_{j=1}^{p} s_j = 0 \text{ et } \sum_{i=1}^{n} e_i = 0$$

Où p désigne une période.
On utilise ce modèle quand, en reliant minima et maxima, on obtient une sorte de cône.

3 Ajustement de la tendance

3.1 Ajustement linéaire

Formule: Méthode des moindres carrés

Elle vient de la recherche des paramètres $a, b \in \mathbb{R}$ minimisant la fonctionnelle suivante :

$$\sum_{i=1}^{n} (y_i - (at_i + b))^2$$

ce qui nous donne :

$$\begin{aligned} \hat{a} &= \frac{\sum_{i=1}^{n}(y_i - \bar{y})(t_i - \bar{t})}{\sum_{i=1}^{n}(t_i - \bar{t})^2} \\ \hat{b} &= \bar{y} - \hat{a}\bar{t} \end{aligned}$$

Formule: Méthode des deux points

Cette méthode consiste à choisir arbitrairement deux points par lesquels on fait passer une droite.
La réalisation de cette méthode se fait en général en prenant deux sous-suites, et en prenant les points médians de chaque sous-série.
Cette méthode s'avère efficace en présence de points aberrants, chose que la méthode des moindres carrés ne prend pas en compte.

Propriété: Appréciation des régression linéaire

Un moyen de qualifier la qualité de la regression linéaire est d'utiliser le coefficiet de corrélation linéaire, noté r, et défini par :

$$r = \frac{\mathrm{cov}(y, t)}{\sigma_y \sigma_t}$$

En effet, en réécrivant l'expression, on peut montrer que :

$$r^2 = \frac{\text{Variance expliquée}}{\text{Variance totale}}$$

3.2 Ajustement polynomial

Formule: Polynôme des moindres carrés

On minimise la même fonction que précédemment, mais en cherchant cette fois non plus a et b d'une régression linéaire mais $a_i,\ i = 0, .., d$ d'un polynôme de degré d. En notant :

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ et } T = \begin{pmatrix} 1 & t_1 & \cdots & t_1^d \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & \cdots & t_n^d \end{pmatrix}$$

On obtient :

$$\theta^{MC} = \begin{pmatrix} a_0 \\ \vdots \\ a_d \end{pmatrix} = ({}^tTT)^{-1} \times {}^tTY$$

3.3 Ajustement non linéaire

On a deux cas :

1. Soit on se ramène à un ajustement linéaire via un changement de variable
2. Soit on cherche à déterminer les coefficients restants via une méthode à d points (d étant le nombre de paramètres à estimer), ou en minimisant le carré des erreurs (ce qui ne donne pas toujours une formule explicite).

4 Lissage par moyenne mobile

Définition: Moyenne mobile simple

On note $MM(k)$ la série des moyennes mobiles d'ordre k de la série $(y_j)_{j=1...n}$, et on a :
— lorsque k est pair et vaut $2m$:

$$MM(k)_j = \frac{y_{j-m+1} + ... + y_j + y_{j+1} + ... + y_{j+m}}{2m}$$

— lorsque k est impair et vaut $2m+1$:

$$MM(k)_j = \frac{y_{j-m} + ... + y_j + y_{j+1} + ... + y_{j+m}}{2m+1}$$

pour $j = m+1, ..., n-m$.

Définition: Moyenne mobile centrée

La série notée $MMC(k)$ uniquement pour k pair et définie par :

$$MMC(k)_j = \frac{MM(k)_{j-1} + MM(k)_j}{2}$$

Propriété:

— La série $MM(p)$ ou $MMC(p)$ ne possède plus de composante saisonnière de période p.
— Une moyenne mobile atténue l'aplitude des fluctuations irrégulières d'une chronique.

5 Décomposition d'une série chronologique

Formule: Étapes de la décomposition

1. La désaisonnalisation
   (a) Lissage par moyennes mobiles : on construit la série des moyennes mobiles d'ordre p, la saisonnalité (centrées si p pair).
   (b) Constrution de la série des différences / quotients : observation - série des moyennes mobiles ou obs / MM
   (c) Calcul des coefficients saisonniers non centrés : moyennes des différences pour chaque saison
   (d) Centrage des coefficients saisonniers : moyennes des p coefficients non centrés, puis on centre les coefficients saisonniers.
   (e) Constrution de la série corrigée des variations saisonnières : observation - composante saisonnière (selon, bien sûr, la saison) ou division.
2. La série lissée des prévisions
   (a) Ajustement d'une tendance : regression linéaire (ou autre) sur la CVS
   (b) Construction de la série lissée des prévisions : résultat de la régression + coefficient saisonnier. $= \hat{y}_i$ (ou $= \hat{f}_i(1 + \hat{s}_i)$)