Ajustement de courbes de tendance par des méthodes de régression *

Jean-Marie Dufour †
Université de Montréal

Première version: Février 1987
Révisions: Mars 2002
Cette version: 10 février 2003
Compilé: 10 février 2003, 10:38am

*Cette recherche a bénéficié du support financier de la Chaire de recherche du Canada en économétrie, du Conseil des Arts du Canada (Bourse Killam), du Conseil de recherche en sciences humaines du Canada, du Conseil de recherche en sciences naturelles et en génie du Canada, de la Fondation Alexander von Humboldt (Allemagne), de l'Institut de Finance mathématique de de Montréal (IFM2), du Réseau canadien de centres d'excellence (projet MITACS) et du Fonds FCAR du Québec.

† L'auteur est titulaire de la Chaire de recherche du Canada en économétrie. Centre interuniversitaire de recherche en analyse des organisations (CIRANO), Centre interuniversitaire de recherche en économie quantitative (CIREQ) et Département de sciences économiques, Université de Montréal. Adresse postale: Département de sciences économiques, Université de Montréal, C.P. 6128 succursale Centre Ville, Montréal, Québec, Canada H3C 3J7. TEL: (514) 343 2400; FAX: (514) 343 5831; courriel: jean.marie.dufour@umontreal.ca. Page Web: http://www.fas.umontreal.ca/SCECO/Dufour .

Table des matières

Liste de définitions, propositions et théorèmes **i**

1. Modèles additifs **1**

2. Modèle polynomial saisonnier **2**

3. Estimation **3**

4. Désaisonnalisation **4**

5. Notes bibliographiques **5**

Liste de définitions, propositions et théorèmes

1. Modèles additifs

Un modèle fréquemment utilisé pour l'analyse de la tendance dans une série chronologique consiste à supposer que

$$Y_t = f(t) + u_t\,, \quad t = 1, \ldots, n$$

où u_t est une perturbation aléatoire avec $\mathsf{E}(u_t) = 0$.Voici quelques formes fréquemment utilisées pour $f(t)$.

1. $f(t) = \beta_0 + \beta_1 t$ — Tendance linéaire
2. $f(t) = \beta_0 + \beta_1 t + \beta_2 t^2$ — Courbe parabolique
3. $f(t) = \beta_0 + \beta_1 t + \cdots + \beta_k t^k$ — Tendance polynominale
4. $f(t) = \beta_0 + \beta_1 r^t$ — Courbe exponentielle
5. $f(t) = \dfrac{1}{\beta_0 + \beta_1 r^t}\,, \quad$ où $r > 0$ — Courbe logistique
6. $f(t) = \exp\{\beta_0 + \beta_1 r^t\}\,, \quad$ où $r > 0$ — Courbe de Gomperz
7. $f(t) = \beta_0 + \beta_1 t + \sum_{j=1}^{4} \gamma_j S_t^j$ — Modèle trimestriel de Buys-Ballot

   où $\sum_{j=1}^{\ell} \gamma_j = 0$ et

$$\begin{aligned} S_t^j &= 1,\ \text{si } t \text{ est une observation du trimestre } j\,, \\ &= 0,\ \text{autrement}\,. \end{aligned}$$

8. $f(t) = \beta_0 + \sum_{i=1}^{k} Z_t^i \beta_i + \sum_{j=1}^{\ell} S_t^j \gamma_j\,,$ — Modèle linéaire saisonnier

   où Z_t^j est une fonction connue de t et $\sum_{j=1}^{\ell} \gamma_j = 0$.

On peut ramener certains modèles à une forme additive au moyen d'une transformation :

$$h(Y_t, Y_{t-1}, \ldots, Y_{t-p}) = f(t) + u_t\,.$$

Par exemple,

$$Y_t = e^{f(t)} u_t^* \tag{1.1}$$

où u_t^* est une perturbation aléatoire telle que $u_t^* > 0$. Il s'ensuit que

$$\log(Y_t) = f(t) + u_t\,,$$

et

$$Y_t = \frac{1}{1+\exp\{f(t)+u_t\}}\,, \qquad \text{Modèle logistique généralisé} \tag{1.2}$$

$$1 - Y_t = \frac{\exp\{f(t)+u_t\}}{1+\exp\{f(t)+u_t\}}\,, \tag{1.3}$$

$$\frac{1-Y_t}{Y_t} = \exp\{f(t)+u_t\}\ , \tag{1.4}$$

$$\log\left(\frac{1-Y_t}{Y_t}\right) = f(t) + u_t\,. \tag{1.5}$$

2. Modèle polynomial saisonnier

Le modèle

$$Y_t = \beta_0 + \beta_1 t + \ldots + \beta_k t^k + \sum_{j=1}^{\ell} S_t^j \gamma_j + u_t\ ,$$

avec $\mathsf{E}(u_t) = 0$ est particulièrement utile en pratique. Les coefficients γ_j peuvent s'interpréter de la manière suivante. Si

$$\delta_j \equiv \text{constante lorsqu'une observation provient de la saison } j\ ,$$

où $j = 1, \ldots, \ell$, et

$$\begin{aligned} \ell &= 4 \quad \text{pour données trimestrielles,} \\ \ell &= 12 \quad \text{pour des données mensuelles,} \end{aligned}$$

alors

$$\delta_j = \beta_0 + \gamma_j\,.$$

Il faut $\ell+1$ coefficients pour représenter ℓ constantes différentes. Il est important d'imposer une contrainte sur les γ_j pour assurer que tous les coefficients soient identifiés. Habituellement, on suppose :

$$\sum_{j=1}^{\ell} \gamma_j = 0\,.$$

Dans ce cas,

$$\beta_0 = \frac{1}{\ell}\sum_{j=1}^{\ell}\delta_j = \text{constante moyenne entre les saisons}$$

$$
\begin{aligned}
\gamma_j &= \delta_j - \beta_0 = \delta_j - \frac{1}{\ell}\sum_{j=1}^{\ell}\delta_j \\
&= \text{déviation de } \delta_j \text{ par rapport à la valeur moyenne.}
\end{aligned}
$$

On pourrait aussi imposer :

$$\gamma_1 = 0$$

d'où

$$
\begin{aligned}
\delta_1 &= \beta_0 \, , \\
\delta_j &= \beta_0 + \gamma_j, \quad j = 1, \ldots, \ell - 1 \, ,
\end{aligned}
$$

et

$$
\begin{aligned}
\gamma_j &= \delta_j - \delta_1, \quad j = 2, \ldots, \ell \\
&= \text{déviation de la constante de la saison } j \\
&\quad\ \text{par rapport à la constante de la saison } 1.
\end{aligned}
$$

Toutefois, c'est une paramétrisation habituellement moins intéressante.

Si on désire estimer le modèle (2.1) avec la contrainte

$$\sum_{j=1}^{\ell}\gamma_j = 0 \ \Rightarrow \gamma_\ell = -\sum_{j=1}^{\ell-1}\gamma_j \, ,$$

on peut écrire

$$Y_t = \beta_0 + \beta_1 t + \cdots + \beta_k t^k + \sum_{j=1}^{\ell-1}\gamma_j \left(S_t^j - S_t^\ell\right) + u_t$$

et estimer $\beta_0, \gamma_1, \ldots, \gamma_{\ell-1}$ ainsi que

$$\gamma_\ell = -\sum_{j=1}^{\ell-1}\gamma_j \, .$$

3. Estimation

Si on suppose que

$$
\begin{aligned}
\mathsf{E}(u_t) = 0, \quad \mathsf{E}(u_s u_t) &= \sigma^2, \quad \text{si } s = t \\
&= 0, \quad \text{si } s \neq t
\end{aligned}
$$

il est naturel de chercher à estimer l'équation

$$Y_t = f(t) + u_t$$

par la méthode des moindres carrés (MC).

Si

$$Y_t = \beta_0 + \sum_{i=1}^{k} Z_t^i \beta_i + \sum_{j=1}^{\ell} S_t^j \gamma_j + u_t , \quad t = 1, \ldots , n$$

avec

$$\sum_{j=1}^{\ell} \gamma_j = 0 .$$

En particulier,

$$\begin{aligned} Y_t &= \beta_0 + \sum_{i=1}^{k} \beta_i t^i + \sum_{j=1}^{\ell} S_t^j \gamma_j + u_t \\ &= \beta_0 + \sum_{i=1}^{k} \beta_i t^i + \sum_{j=1}^{\ell-1} \left(S_t^j - S_t^\ell \right) \gamma_j + u_t, \quad t = 1, \ldots , n , \end{aligned}$$

on peut employer les moindres carrés ordinaires (MCO). Si

$$u_t \overset{i.i.d.}{\sim} N\left[0, \sigma^2\right] , \quad t = 1, \ldots , n ,$$

on peut se servir des tests et IC's usuels.

Si $g(t)$ est une fonction non linéaire, on peut utiliser les moindres carrés non linéaires.

4. Désaisonnalisation

Si une série peut s'écrire

$$Y_t = X_t + S_t + u_t$$

où S_t représente les variations saisonnières, on obtient la série désaisonnalisée correspondante en calculant

$$Y_t^{DS} = Y_t - S_t = X_t + u_t .$$

Par exemple, dans (2.1), nous avons :

$$X_t = \beta_0 + \beta_1 t + ... + \beta_k t^k ,$$

$$S_t \;=\; \sum_{j=1}^{\ell} S_t^j \gamma_j \;,$$

d'où

$$Y_t^{DS} = Y_t - \sum_{j=1}^{\ell} S_t^j \gamma_j \;.$$

Si les γ_j sont estimées, on calcule

$$Y_t^{DS} = Y_t - \sum_{j=1}^{\ell} S_t^j \hat{\gamma}_j \;.$$

5. Notes bibliographiques

Voir Gouriéroux et Monfort (1983, Chap. II), Gouriéroux et Monfort (1990, Chap. II) et Diebold (1998, Chapter 4).

Références

Diebold, F. (1998), *Elements of Forecasting*, South-Western College Publishing, Cincinnati (Ohio).

Gouriéroux, C. et Monfort, A. (1983), *Cours de séries temporelles*, Economica, Paris.

Gouriéroux, C. et Monfort, A. (1990), *Séries temporelles et modèles dynamiques*, Economica, Paris.