

Série temporelles: TD 1

Quentin Moret

Louis Martin

quentin.moret@student.ecp.fr

louis.martin@student.ecp.fr

February 19, 2017

Exercice 1 : Le Fed model

Partie 1 : Estimation du modèle

Question 1

Il y a clairement une relation entre les deux séries. Cependant la relation linéaire ne semble pas si évidente que cela. On a tout de même intérêt à essayer un modèle linéaire en premier afin d'avoir une baseline avant de tester des modèles plus compliqués.

La densité empirique de l'endogène n'est pas gaussienne, elle a plus l'allure d'un mélange de plusieurs gaussiennes. Cela n'est pas problématique pour une régression linéaire.

Pour une méthode de l'estimation linéaire du premier ordre cela n'est pas problématique.

Question 2

On veut trouver une relation du type $y = \alpha + \beta x$ La méthode des moindres carrés ordinaires consiste à trouver l'estimateur qui minimise l'erreur au carré, i.e.:

$$(\hat{\alpha}, \hat{\beta}) = \arg \min_{\alpha, \beta} \sum (y_i - \alpha - \beta x_i)^2$$

En dérivant par rapport à α on obtient

$$\sum -2(y_i - \hat{\alpha} - \hat{\beta}x_i) = 0 \iff \bar{y} = \hat{\alpha} + \hat{\beta}\bar{x}$$

En dérivant par rapport à $\hat{\beta}$ on obtient

$$\sum -2x_i(y_i - \hat{\alpha} - \hat{\beta}x_i) = 0 \iff \sum x_i(y_i - \bar{y} + \hat{\beta}\bar{x} - \hat{\beta}x_i) = 0$$

$$\iff n \sum x_i y_i - (\sum x_i)(\sum y_i) + \hat{\beta}(\sum x_i)^2 - \hat{\beta}n \sum x_i^2 = 0$$

$$\iff \hat{\beta} = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{n \sum x_i^2 - (\sum x_i)^2}$$

$$\iff \hat{\beta} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\iff \hat{\beta} = \frac{Cov(x, y)}{Var(x)} \text{ (covariance et variance empiriques)}$$

On retrouve donc bien les formules demandées avec $x = r$ et $y = \frac{E}{P}$.

Question 3

L'estimateur des MCO est sans biais ($E[\hat{\beta}] = \beta$), consistant ($\lim_{T \rightarrow \infty} \text{Var}[\hat{\beta}] = 0$) et c'est le meilleur estimateur parmi les estimateurs linéaires.

Démontrons la consistance.

$$\begin{aligned}\hat{\beta} &= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \\ &\xrightarrow{T \rightarrow \infty} \frac{\text{Cov}(x, y)}{\text{Var}(x)} \text{ (d'après la loi des grands nombres)} \\ &\xrightarrow{T \rightarrow \infty} \frac{\text{Cov}(x, \alpha + \beta x + \epsilon)}{\text{Var}(x)} \\ &\xrightarrow{T \rightarrow \infty} \frac{\text{Cov}(x, \beta x) + \text{Cov}(x, \epsilon)}{\text{Var}(x)} \\ &\xrightarrow{T \rightarrow \infty} \beta + \frac{\text{Cov}(x, \epsilon)}{\text{Var}(x)} \\ &\xrightarrow{T \rightarrow \infty} \beta \text{ (Parce que } x \text{ est indépendant de } \epsilon)\end{aligned}$$

Donc

$$\lim_{T \rightarrow \infty} \text{Var}(\hat{\beta}) = \lim_{T \rightarrow \infty} E(\hat{\beta} - \beta) = 0$$

Question 4

Insérer figure ici

Le coefficient associé à $TX10$ est positif et supérieur à 1. Les statistiques de Student indiquent une grande significativité des coefficients ($pvalue < 2 * 10^{-16}$). La statistique de Fisher indique que la non nullité simultanée des coefficients est statistiquement significative ($pvalue < 2 * 10^{-16}$ de même).

En revanche le R^2 est très faible, ce qui n'est pas surprenant au vu de la variabilité du graphique opposant PER à $TX10$.

Montrons que $R^2 = \frac{\text{cov}(X_t, Y_t)^2}{\text{Var}(X_t)\text{Var}(Y_t)}$

Nous avons:

$$R^2 = \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (y_i - \bar{y})^2}$$

En utilisant le fait que $\hat{y}_i = \hat{\beta}x_i + \hat{\alpha}$ et que $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$,

$$\begin{aligned}\sum_i (\hat{y}_i - \bar{y})^2 &= \sum_i (\hat{\alpha} + \hat{\beta}x_i - \bar{y})^2 \\ \sum_i (\hat{y}_i - \bar{y})^2 &= \sum_i (\bar{y} - \hat{\beta}\bar{x} + \hat{\beta}x_i - \bar{y})^2 \\ \sum_i (\hat{y}_i - \bar{y})^2 &= \hat{\beta}^2 \sum_i (x_i - \bar{x})^2\end{aligned}$$

En utilisant le le fait que $\hat{\beta} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}$, on obtient

$$\sum_i (\hat{y}_i - \bar{y})^2 = \frac{(\sum_i (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum_i (x_i - \bar{x})^2}$$

Enfin,

$$R^2 = \frac{(\sum_i (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}$$

$$R^2 = \frac{cov(X, Y)^2}{\sigma_X^2 \sigma_Y^2}$$

Partie 2 : Estimation d'une nouvelle spécification et comparaison

Question 5

En utilisant le taux d'intérêt réel comme variable prédictive, on obtient une régression au pouvoir prédictif amélioré : le R^2 évolue de 0.272 à 0.312

Question 6

Les mesures d'erreur RMSE et MAE confirment le précédent résultat. La RMSE passe de 4.12 à 4.09 en utilisant les taux réels, et la MAE passe de 3.36 à 3.26.

Question 7

En utilisant la routine développée pour évaluer savoir si le modèle avec les taux réels est plus pertinent que l'autre, on obtient une statistique très grande quelque soit la valeur de h (en moyenne $|SDM| \approx 8 \times 10^6$, à comparer avec 1.96). On en conclut que le modèle utilisant les taux réels apporte une réelle amélioration.

Question 8.a

Les β sont très instables, et les intervalles de confiance très grands. Le modèle est donc imparfait. Nous allons étudier à la question suivante si ces instabilités sont aléatoires ou sont liés à une évolution structurelle de la relation entre les données.

Question 8.b

Le test CUSUM permet d'évaluer la stabilité du coefficient β . On évalue une statistique dont l'hypothèse nulle est la constance du coefficient. On calcule d'abord les résidus par estimation récursive du coefficient. On compare ensuite l'évolution de la somme de ces résidus à une marche aléatoire (de variance \sqrt{n}). Si la somme des résidus ne dépasse pas un certain seuil, alors on peut affirmer qu'ils sont aléatoires et donc que le coefficient est stable (pas d'évolution structurelle du modèle au cours du temps).

On constate en effet sur le test CUSUM sur les résidus récurrents que le seuil est largement dépassé, notre modèle n'est donc pas stable et n'est pas adapté.

Question 9

La lente décroissance des autocorrélations semble indiquer une tendance non éliminée. Avec l'autocorrélation partielle en revanche, on ne note pas d'autocorrélation avec lag.