

# Autocorrélation des erreurs

GSF-6053

Hiver 2025

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Modèle de Base avec Autocorrélation</b>	<b>2</b>
2.1	Matrice de Variance-Covariance des Erreurs . . . . .	3
2.2	Preuve de la Forme de $\Omega$ . . . . .	4
<b>3</b>	<b>La Transformation de Prais-Winsten</b>	<b>6</b>
<b>4</b>	<b>La Transformation de Cochran-Orcutt</b>	<b>7</b>
<b>5</b>	<b>Le Problème de <math>\rho</math></b>	<b>7</b>
5.1	Estimateur Usuel ( <a href="#">gujarati2010</a> ) . . . . .	7
5.2	Estimateur de Hildreth-Lu . . . . .	7
<b>6</b>	<b>Diagnostic de l'Autocorrélation</b>	<b>8</b>
6.1	Test de Durbin-Watson . . . . .	8
6.2	Le Test de Breusch-Godfrey . . . . .	10
6.3	Estimateur Robuste de Newey-West . . . . .	10
<b>7</b>	<b>Conclusion</b>	<b>11</b>

# 1 Introduction

[gujarati2010](#), [wooldridge2010](#)

L'autocorrélation ou la corrélation sérielle des erreurs dans les séries chronologiques est assez fréquente. Elle découle souvent du fait qu'il y a une certaine inertie dans les données économiques et financières, c'est-à-dire que les observations passées se reflètent souvent dans les observations présentes et futures. Cela peut amener des problèmes de corrélation sérielle des erreurs si une variable autocorrélée est omise du modèle, par exemple.

Il est important de s'assurer que lorsque vous détectez de l'autocorrélation, elle n'est pas due à la non-stationnarité de  $Y$  et  $X$ . En effet, si ces deux quantités sont non stationnaires, les erreurs le seront possiblement aussi et seront possiblement autocorrélées.

Même si le modèle est bien spécifié, on peut penser que les erreurs (les chocs) affectant les marchés boursiers aujourd'hui ont des chances d'influencer la magnitude et le signe des chocs affectant les marchés boursiers demain aussi. Dans la plupart des cas en finance, l'autocorrélation sera positive, mais il est possible en théorie d'avoir de l'autocorrélation négative.

## 2 Modèle de Base avec Autocorrélation

[gujarati2010](#), [wooldridge2010](#)

Soit le modèle de base :

$$y = X\beta + u \quad (1)$$

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_{K-1} X_{K-1,t} + u_t \quad (2)$$

$$u_t = \rho u_{t-1} + \epsilon_t \quad (3)$$

Où  $\epsilon_t$  est un bruit blanc de moyenne nulle et de variance constante dans le temps.

**Différence avec le modèle linéaire de base :**



La matrice de variance-covariance des erreurs sera affectée, ce qui impliquera, comme dans le cas hétéroscédastique, que l'estimateur OLS standard ne sera pas BLUE et que la variance des estimateurs OLS sera incorrecte, bien que l'estimateur soit sans biais. La démonstration de ceci est du même ressort que celle faite pour l'hétéroscédasticité. En effet, bien que l'autocorrélation implique des erreurs homoscedastiques (la variance est constante), les termes hors diagonale ne sont pas nuls comme dans le cas de base du modèle linéaire.

## 2.1 Matrice de Variance-Covariance des Erreurs

$$\mathbf{E}(uu') = \begin{bmatrix} \mathbb{E}(u_1^2) & \mathbb{E}(u_1u_2) & \mathbb{E}(u_1u_3) & \dots & \mathbb{E}(u_1u_T) \\ \mathbb{E}(u_2u_1) & \mathbb{E}(u_2^2) & \mathbb{E}(u_2u_3) & \dots & \mathbb{E}(u_2u_T) \\ \mathbb{E}(u_3u_1) & \mathbb{E}(u_3u_2) & \mathbb{E}(u_3^2) & \dots & \mathbb{E}(u_3u_T) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}(u_Tu_1) & \mathbb{E}(u_Tu_2) & \mathbb{E}(u_Tu_3) & \dots & \mathbb{E}(u_T^2) \end{bmatrix}$$

Avec :

$$\mathbf{E}(uu') = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{bmatrix} = \sigma^2 \Omega$$

Où :

$$\Omega = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{bmatrix}$$



## 2.2 Preuve de la Forme de $\Omega$

### Hypothèses de départ :

1.  $\epsilon_t$  est un bruit blanc, c'est-à-dire :

$$\mathbb{E}(\epsilon_t) = 0, \quad \mathbb{V}(\epsilon_t) = \sigma^2, \quad \forall t \quad \text{et} \quad \mathbb{E}(\epsilon_t \epsilon_s) = 0 \quad \forall t \neq s$$

2.  $|\rho| < 1$  (absence de racine unitaire).

Ceci implique que  $u_t$  est un processus stationnaire, c'est-à-dire :

$$(a) \quad \mathbb{E}(u_t) = 0$$

$$(b) \quad \mathbb{V}(u_t) = \frac{\sigma^2}{1-\rho^2}, \quad \forall t$$

$$(c) \quad \mathbb{E}(u_t u_{t-j}) = \frac{\rho^j \sigma^2}{1-\rho^2} \quad \forall j \geq 1$$

3.  $\epsilon_t$  n'est pas corrélé avec les  $u_s$  dont l'indice  $s$  est antérieur à  $t$ , c'est-à-dire :

$$\epsilon_t \perp u_{t-s}, \quad \forall s \geq 1$$

### Développement :

$$\begin{aligned} u_t &= \rho u_{t-1} + \epsilon_t \\ &= \rho(\rho u_{t-2} + \epsilon_{t-1}) + \epsilon_t \\ &= \rho^2 u_{t-2} + \rho \epsilon_{t-1} + \epsilon_t \\ &= \dots \\ u_t &= \sum_{s=0}^{\infty} \rho^s \epsilon_{t-s} \end{aligned} \tag{4}$$

### Calcul de la Variance de $u_t$ :



$$\begin{aligned}
\mathbb{V}(u_t) &= \mathbb{V} \left( \sum_{s=0}^{\infty} \rho^s \epsilon_{t-s} \right) \\
&= \sum_{s=0}^{\infty} (\rho^s)^2 \mathbb{V}(\epsilon_{t-s}) \\
&= \sigma^2 \sum_{s=0}^{\infty} \rho^{2s} \\
&= \sigma^2 \left( \frac{1}{1 - \rho^2} \right) \quad (\text{série géométrique convergente pour } |\rho| < 1) \quad (5)
\end{aligned}$$

Ainsi :

$$\mathbb{V}(u_t) = \frac{\sigma^2}{1 - \rho^2}$$

**Calcul des Covariances :**

$$\begin{aligned}
\mathbb{E}(u_t u_{t-j}) &= \mathbb{E} \left( \sum_{s=0}^{\infty} \rho^s \epsilon_{t-s} \sum_{k=0}^{\infty} \rho^k \epsilon_{t-j-k} \right) \\
&= \sum_{s=0}^{\infty} \sum_{k=0}^{\infty} \rho^{s+k} \mathbb{E}(\epsilon_{t-s} \epsilon_{t-j-k}) \\
&= \sum_{s=0}^{\infty} \rho^{s+j+s} \mathbb{E}(\epsilon_{t-s} \epsilon_{t-j-s}) \\
&= \sum_{s=0}^{\infty} \rho^{2s+j} \mathbb{E}(\epsilon_{t-s} \epsilon_{t-j-s}) \\
&= \rho^j \sigma^2 \sum_{s=0}^{\infty} \rho^{2s} \\
&= \rho^j \sigma^2 \left( \frac{1}{1 - \rho^2} \right) \\
&= \frac{\rho^j \sigma^2}{1 - \rho^2} \quad (6)
\end{aligned}$$

Ainsi, pour toute autocorrélation d'ordre  $j$ , on a :



$$\mathbb{E}(u_t u_{t-j}) = \frac{\rho^j \sigma^2}{1 - \rho^2}$$

### 3 La Transformation de Prais-Winsten

[gujarati2010](#), [wooldridge2010](#)

Comme on l'a fait dans le cas de l'hétéroscédasticité, on peut utiliser un modèle transformé (donc un GLS ou un FGLS) pour utiliser les OLS en présence d'autocorrélation des erreurs. La transformation de Prais-Winsten consiste à multiplier les éléments de la régression par la matrice  $P^{-1}$ .

$$P^{-1} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_{T-1} \\ Y_T \end{bmatrix} = \begin{bmatrix} \sqrt{1 - \rho^2} Y_1 \\ -\rho Y_1 + Y_2 \\ -\rho Y_2 + Y_3 \\ \vdots \\ -\rho Y_{T-1} + Y_T \end{bmatrix}$$

La transformation sera la même pour chaque régresseur  $X$ . Le modèle transformé s'écrit :

$$\sqrt{1 - \rho^2} Y_1 = \sqrt{1 - \rho^2} \beta_0 + \beta_1 \sqrt{1 - \rho^2} X_{11} + \beta_2 \sqrt{1 - \rho^2} X_{21} + \dots + \sqrt{1 - \rho^2} u_1 \quad (7)$$

$$\begin{aligned} Y_t - \rho Y_{t-1} &= (1 - \rho) \beta_0 + \beta_1 (X_{1t} - \rho X_{1,t-1}) + \beta_2 (X_{2t} - \rho X_{2,t-1}) + \dots \\ &\quad + \beta_{K-1} (X_{K-1,t} - \rho X_{K-1,t-1}) + u_t - \rho u_{t-1} \end{aligned} \quad (8)$$

Il faut porter attention à la constante dans ce modèle qui n'est plus la constante d'origine. Si  $\rho$  est inconnu, ce modèle n'est plus linéaire, mais on peut toujours l'estimer par OLS.



## 4 La Transformation de Cochran-Orcutt

[gujarati2010](#), [wooldridge2010](#)

Une autre façon de transformer le modèle est de laisser tomber la première observation et de transformer les  $Y$  comme suit :

$$Y_t - \rho Y_{t-1}, \quad t = 2, \dots, T$$

Les  $X$  sont transformés de la même façon.

Le modèle transformé de Cochran-Orcutt s'écrit de la même façon que Prais-Winsten, mais sans la première observation :

$$\begin{aligned} Y_t - \rho Y_{t-1} = & (1 - \rho)\beta_0 + \beta_1(X_{1t} - \rho X_{1,t-1}) + \beta_2(X_{2t} - \rho X_{2,t-1}) + \dots \\ & + \beta_{K-1}(X_{K-1,t} - \rho X_{K-1,t-1}) + u_t - \rho u_{t-1} \end{aligned} \quad (9)$$

## 5 Le Problème de $\rho$

$\rho$  est inconnu. Il faut donc trouver un estimateur. Plusieurs sont suggérés dans la littérature économétrique.

### 5.1 Estimateur Usuel ([gujarati2010](#))

$$\hat{\rho} = \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^T \hat{u}_{t-1}^2}$$

Où  $\hat{u}_t$  est le résidu OLS de la régression de  $Y$  sur  $X$ . On peut montrer que PLIM  $\hat{\rho} = \rho$ .

### 5.2 Estimateur de Hildreth-Lu

Cet estimateur ne passe pas par les OLS. En balayant l'intervalle  $(-1, 1)$ , on peut choisir la valeur de  $\rho$  qui minimise la somme des carrés des erreurs du modèle transformé. Les étapes sont :



1. Choisir  $\rho^{(1)} = -0.99999$ .
2. Obtenir le modèle transformé correspondant.
3. Appliquer les OLS et sauvegarder  $\hat{u}'\hat{u}_{\rho^{(1)}}$ .
4. Choisir une nouvelle valeur :  $\rho^{(2)} = \rho^{(1)} + \text{step}$ .
5. Obtenir le modèle transformé correspondant.
6. Appliquer les OLS et sauvegarder  $\hat{u}'\hat{u}_{\rho^{(2)}}$ .
7. Continuer jusqu'à ce que  $(-1, 1)$  soit couvert.
8. Choisir la valeur  $\rho^{(s)}$  telle que

$$\hat{u}'\hat{u}_{\rho^{(s)}} = \min [\hat{u}'\hat{u}_{\rho^{(i)}}]_i$$

L'estimateur FGLS obtenu à partir d'un estimateur de  $\rho$  sera convergent. Donc, lorsque la matrice d'information  $P$  dépend de paramètres inconnus qu'il faut estimer, on perd la propriété BLUE. L'estimateur qui est un FGLS au lieu de GLS reste convergent si l'estimateur des paramètres inconnus sur lequel il est fondé est convergent.

## 6 Diagnostic de l'Autocorrélation

[gujarati2010](#)

On peut suspecter l'autocorrélation en faisant un graphique entre les résidus aujourd'hui et hier. Ce n'est pas un test statistique. Il existe des tests d'autocorrélation des erreurs.

### 6.1 Test de Durbin-Watson

C'est un test à bornes. Vous avez deux points critiques à regarder dans la table. À l'origine, ce test était à bornes, car on ne connaissait pas les vrais points critiques seulement des bornes sous l'hypothèse de normalité. Maintenant, on a trouvé les vrais points critiques.

On teste :





$$H_0 : \rho = 0 \quad \text{contre} \quad H_A : \rho \neq 0$$

La statistique de test est :

$$d = \frac{\sum_{t=2}^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^T \hat{u}_t^2} \approx 2(1 - \hat{\rho}) \quad (10)$$

### Hypothèses du test de Durbin-Watson :

1. Le modèle de régression inclut une constante pour pouvoir calculer RSS.
2. Les variables explicatives sont non stochastiques.
3. L'autocorrélation est d'ordre 1. On ne peut pas utiliser Durbin-Watson avec des ordres supérieurs.
4. On ne peut pas utiliser ce test dans le cas de modèles autorégressifs.
5. Le test n'acommode pas les observations manquantes.

Pour faire le test de Durbin-Watson, il faut effectuer une régression OLS et obtenir les résidus. Ensuite, on calcule la statistique de Durbin-Watson avec ces résidus. On compare la statistique avec les points critiques de la distribution pour les statistiques de Durbin-Watson pour une taille d'échantillon donnée et un  $K$  donné. On suit la règle de décision suivante :

$$\left\{ \begin{array}{ll} d < d_L & \rightarrow \text{rejeter } H_0 \\ d_L < d < d_U & \rightarrow \text{test non concluant} \\ d_U < d < 4 - d_U & \rightarrow \text{ne pas rejeter } H_0 \text{ et } H_0^* \\ 4 - d_U < d < 4 - d_L & \rightarrow \text{test non concluant} \\ d > 4 - d_L & \rightarrow \text{rejeter } H_0^* \end{array} \right.$$

La statistique de Durbin-Watson est une « institution » en économétrie parce qu'il s'agit d'un des premiers tests à bornes dérivés, mais les hypothèses sous-jacentes sont assez contraignantes.



## 6.2 Le Test de Breusch-Godfrey

Ce test est plus général et plus populaire que Durbin-Watson. Il permet de tester pour des formes d'autocorrélation avec plusieurs retards ( $q$ ). L'intérêt du test est qu'il se généralise facilement pour  $q$  lags. Pour  $q$  retards, la régression artificielle est de  $\hat{u}_t$  sur une constante,  $X_{1t}, \dots, X_{kt}$  et  $\hat{u}_{t-1}, \dots, \hat{u}_{t-q}$ .

$$\hat{u}_t = \alpha_1 + \alpha_2 X_t + \dots + \rho_1 \hat{u}_{t-1} + \dots + \rho_q \hat{u}_{t-q} + \epsilon_t$$

**Hypothèse nulle :**

$H_0 : \rho_1 = \rho_2 = \dots = \rho_q = 0$  contre l'alternative qu'au moins une des autocorrélations

La statistique de test est :

$$BG = (T - q)R^2 \sim \chi^2(q)$$

## 6.3 Estimateur Robuste de Newey-West

[gujarati2010](#), [greene2018](#)

Il existe une forme de correcteur robuste pour les problèmes d'autocorrélation des erreurs qui est une alternative au FGLS lorsque le nombre d'observations est assez élevé. En effet, cette correction est asymptotique. C'est l'estimateur de la variance de Newey-West. L'intuition est la même que le correcteur de White, mais appliquée aux problèmes d'autocorrélation des erreurs. En fait, le charme de l'estimateur de Newey-West est qu'il corrige pour l'autocorrélation ET l'hétéroscédasticité des erreurs (c'est ce qu'on appelle l'estimateur « HAC 1 »). C'est un correcteur très populaire en finance.

Plus généralement, la classe d'estimateurs corrige les déviations dans la matrice de variance-covariance des erreurs par rapport au cas simple.

Nous n'allons pas faire la preuve ou la dérivation, car elle est assez avancée. Pour ceux qui veulent aller plus loin, elle est présentée dans la section 20.5.2 de [greene2018](#). Le correcteur de Newey-West est déjà programmé dans les logiciels d'analyse de données comme Stata.



## 7 Conclusion

L'autocorrélation des erreurs est une problématique courante dans les modèles de régression linéaire appliqués aux séries chronologiques. Les tests diagnostiques tels que Durbin-Watson et Breusch-Godfrey permettent de détecter sa présence et de choisir les méthodes appropriées pour la corriger, assurant ainsi la fiabilité des estimations des paramètres du modèle. Des méthodes de correction comme Prais-Winsten, Cochran-Orcutt et l'estimateur de Newey-West offrent des solutions robustes pour traiter l'autocorrélation et améliorer les performances des modèles économétriques.

