

Introduction à l'économétrie

Méthode des moindres carrés ordinaires en présence d'hétéroscédasticité ou d'autocorrélation des erreurs – Leçon 2

Ce cours vous est proposé par Olivier Baron, Maître de conférences, Université de Bordeaux et par AUNEGe, l'Université Numérique en Économie Gestion.

Table des matières

Préambule	3
Introduction	3
Autocorrélation des perturbations	4
Détection de l'autocorrélation des résidus estimés	4
Détection visuelle.....	4
Le test de Geary.....	5
Le test de Durbin-Watson.....	8
Le test de Breusch – Godfrey	11
Conséquences de l'autocorrélation des erreurs sur la méthode des MCO	12
Estimation d'un modèle en présence d'autocorrélation des perturbations	13
Estimation de ρ par régression des résidus estimés du modèle initial ε_i sur leurs valeurs retardées $\varepsilon_i - 1$	14
Estimation de ρ à partir de la statistique de Durbin-Watson, DW	14
Méthode de Durbin.....	14
Méthode de Cochrane-Orcutt.....	15
Méthode du balayage de Hildreth-Lu	16

Références..... 18

Préambule

Objectifs :

- Comprendre et étudier les conséquences de la non vérification de l'hypothèse H_4 sur la méthode d'estimation par moindres carrés.

Introduction

L'hypothèse H_4 présentée au cours du premier chapitre du cours suppose que la matrice de variances – covariances des perturbations est une matrice scalaire, soit :

$$\text{Var}_{(n,n)} \varepsilon = E \left(\begin{matrix} \varepsilon \\ \varepsilon' \end{matrix} \right) = \sigma^2 \cdot I_n \Leftrightarrow \begin{cases} E(\varepsilon_i^2) = \text{Var}(\varepsilon_i) = \sigma^2 & \forall i = 1, \dots, n \\ E(\varepsilon_i \varepsilon_j) = \text{Cov}(\varepsilon_i, \varepsilon_j) = 0 & \forall i, \forall j, i \neq j \end{cases}$$

Selon cette hypothèse, la variance des perturbations est constante et égale à σ^2 (les erreurs sont homoscédastiques) et les perturbations ne sont pas corrélées entre elles (absence d'autocorrélation des erreurs).

Nous allons nous interroger dans ce chapitre sur les conséquences de la non vérification de cette hypothèse sur la méthode d'estimation par moindres carrés. Clairement, deux situations peuvent conduire à la remise en cause de l'hypothèse H_4 :

- Soit la variance des perturbations n'est pas constante (situation qualifiée d'hétéroscédasticité et plus couramment observée sur des données en coupe transversale) ;
- Soit les erreurs sont corrélées entre elles (situation qualifiée d'autocorrélation et plus souvent observée sur des séries temporelles ou longitudinales).

En pratique, les deux situations peuvent pourtant être simultanément rencontrées. Il faudra toujours tester la présence d'autocorrélation avant l'hétéroscédasticité car les tests permettant d'identifier cette dernière sont sensibles à l'autocorrélation alors que les tests d'autocorrélation (Durbin – Watson en particulier) sont robustes à l'hétéroscédasticité.

Autocorrélation des perturbations

On suppose ici une situation où les perturbations sont homoscédastiques ($\text{Var}(\varepsilon_i) = \sigma^2$) mais sont à présent corrélées entre elles, soit :

$$\text{Var}_{(n,n)} \varepsilon = \begin{pmatrix} \sigma^2 & \sigma_{1,2} & \dots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1} & \sigma_{n,2} & \dots & \sigma^2 \end{pmatrix} \neq \sigma^2 \cdot I_n \quad \Leftrightarrow \quad \text{Cov}(\varepsilon_i, \varepsilon_j) = \sigma_{i,j} \neq 0$$

Nous serons en présence d'une autocorrélation des erreurs lorsque les perturbations sont liées par un processus de reproduction : on parle de processus à mémoire. Ce cas est particulièrement fréquent lorsqu'on travaille sur des séries temporelles.

La **détection d'une éventuelle dépendance des erreurs** ne peut s'effectuer qu'**à partir des résidus estimés**, car eux seuls sont connus. Ainsi, si l'on détecte une telle liaison sur ces résidus, on sera en droit de penser que cette liaison existe aussi au niveau des perturbations du modèle.

Détection de l'autocorrélation des résidus estimés

Détection visuelle

L'analyse graphique des résidus estimés permet le plus souvent de détecter un processus de reproduction des erreurs lorsque :

- Les résidus estimés sont pendant plusieurs périodes consécutives soit positifs, soit négatifs. On parle alors d'autocorrélation positive :

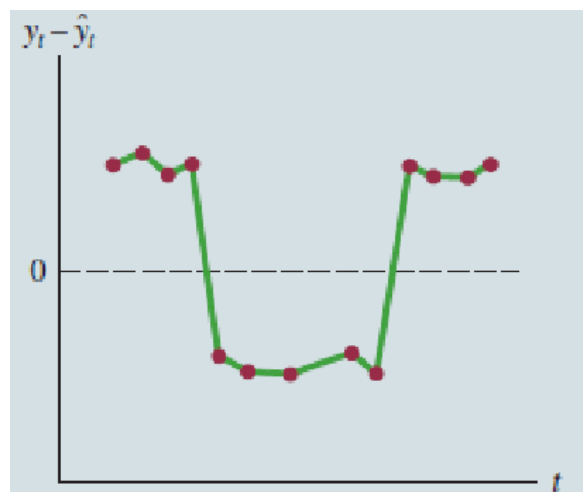


Figure 1 : autocorrélation positive des résidus estimés

- Les résidus estimés sont alternés. On parle alors d'autocorrélation négative :

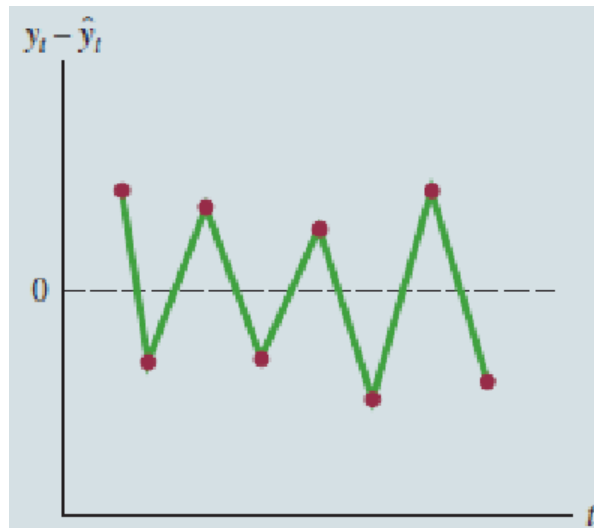


Figure 2 : : autocorrélation négative des résidus estimés

Cependant, le plus souvent, l'analyse graphique des résidus estimés est délicate car les représentations graphiques ne présentent pas de caractéristiques particulières évidentes.

Le test de Geary¹

Ce test non paramétrique est aussi couramment appelé **test des séquences** (« Runs test » en anglais). Le test de Geary se base sur l'observation du graphique des résidus estimés et doit donc **être mené en parallèle de l'analyse visuelle précédente**.

Prenons l'exemple suivant : on estime un modèle linéaire par moindres carrés ordinaires sur un échantillon de $n = 46$ observations, puis on représente graphiquement la série des résidus estimés. Le résultat obtenu est le suivant :

¹ Geary, R.C. « *Relative efficiency of count sign changes for assessing residual autoregression in least squares regression* », *Biometrika*, 57, 1970.

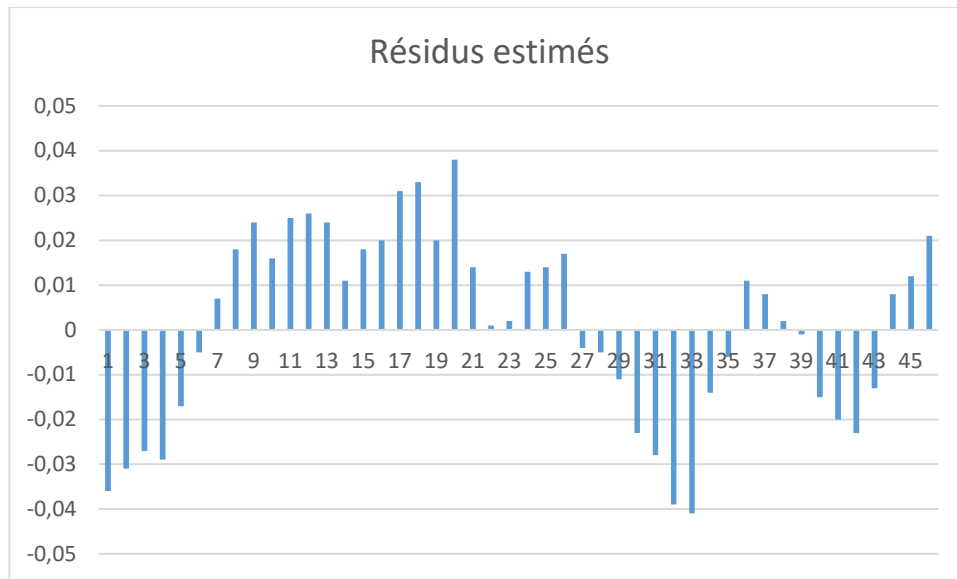


Figure 3 : graphique des résidus estimés

Si l'on examine le graphique, on remarque immédiatement que les résidus obéissent à une structure particulière. Ils sont négatifs au départ, puis deviennent positifs pendant un nombre important de périodes, puis redeviennent négatifs, etc.

Si ces résidus estimés étaient purement aléatoires, devrions-nous observer un tel schéma ? Intuitivement, cela semble peu probable. Cette intuition peut être vérifiée à l'aide du test de Geary ou encore du test des séquences.

On définit une séquence comme une suite ininterrompue de résidus de mêmes signes et la longueur d'une séquence par le nombre d'observations qui la compose. Dans notre exemple, on peut compter 6 séquences dont les longueurs sont :

- Séquence 1 (résidus négatifs) – longueur 6,
- Séquence 2 (résidus positifs) – longueur 20,
- Séquence 3 (résidus négatifs) – longueur 9,
- Séquence 4 (résidus positifs) – longueur 3,
- Séquence 5 (résidus négatifs) – longueur 5,
- Séquence 6 (résidus positifs) – longueur 3.

En examinant comment les séquences apparaîtraient si elles étaient le résultat d'observations strictement aléatoires, on peut construire un test permettant de statuer sur cette situation, c'est-à-dire sur l'indépendance des observations, qui sont ici les résidus estimés de la régression initiale.

Posons $n = n_1 + n_2$, où n_1 est le nombre de résidus négatifs et n_2 le nombre de résidus positifs. Soit S le nombre de séquences observées dans le graphique des résidus estimés.

Sous l'hypothèse nulle selon laquelle les résidus successifs sont indépendants et dès lors que $n_1 > 10$ et $n_2 > 10$, Geary montre que le nombre de séquences S est distribué normalement avec :

$$E(S) = \frac{2n_1n_2}{n} + 1$$

$$Var(S) = \sigma_S^2 = \frac{2n_1n_2(2n_1n_2 - n)}{n^2(n - 1)}$$

Si l'hypothèse nulle d'absence d'autocorrélation des perturbations est vraie, compte tenu des propriétés des distributions gaussiennes on doit donc s'attendre à ce que :

$$\text{Prob}\{E(S) - 1,96 \cdot \sigma_S < S < E(S) + 1,96 \cdot \sigma_S\} = 0,95$$

Décision : On rejette l'hypothèse nulle d'absence d'autocorrélation si le nombre observé de séquences S est en dehors de l'intervalle de confiance précédent (on peut bien sûr choisir un niveau de confiance différent).

Dans l'exemple illustratif on a : $n_1 = 20$; $n_2 = 26$ et $S = 6$. On obtient donc :

$$E(S) = 23,61$$

$$Var(S) = \sigma_S^2 = 10,86$$

ce qui conduit à l'intervalle de confiance suivant :

$$\text{Prob}\{23,61 - 1,96 \cdot \sqrt{10,86} < S < 23,61 + 1,96 \cdot \sqrt{10,86}\} = 0,95$$

soit :

$$\text{Prob}\{17,15 < S < 30,07\} = 0,95$$

Il est clair que cet intervalle ne contient pas la valeur observée $S = 6$. On peut donc **conclure à une autocorrélation des perturbations, qui plus est, positive**. Le nombre de séquences est anormalement faible. De façon symétrique, on conclurait à une autocorrélation négative des perturbations si le nombre de séquences était élevé et dépassait la borne supérieure de l'intervalle.

Le test de Durbin-Watson²

Le test de Durbin-Watson permet de détecter une autocorrélation des erreurs d'ordre 1 de la forme suivante :

$$\varepsilon_i = \rho \cdot \varepsilon_{i-1} + u_i$$

où u_i est une perturbation gaussienne avec des propriétés classiques. Le coefficient ρ est appelé coefficient d'autocorrélation et on dira, dans ce cas, que les erreurs suivent une spécification autorégressive d'ordre 1 (ou encore AR(1)).

On se propose de réaliser le test suivant :

$$H_0 : \rho = 0 \\ \text{contre } \bar{H} : \rho \neq 0$$

Pour cela on construit la statistique de Durbin-Watson définie par :

$$DW = \frac{\sum_{i=2}^n (\hat{\varepsilon}_i - \hat{\varepsilon}_{i-1})^2}{\sum_{i=1}^n \hat{\varepsilon}_i^2}$$

où $\hat{\varepsilon}_i$ sont les résidus estimés du modèle.

De par sa construction, la statistique DW varie entre 0 et 4, et il est simple de montrer que $DW \approx 2(1 - \hat{\rho})$ où $\hat{\rho}$ est l'estimateur de ρ obtenu à partir de la régression des résidus $\hat{\varepsilon}_i$ sur leur valeur passée $\hat{\varepsilon}_{i-1}$:

$$\hat{\rho} = \frac{\text{Cov}(\hat{\varepsilon}_i, \hat{\varepsilon}_{i-1})}{\text{Var } \hat{\varepsilon}_i} = \frac{\sum_{i=2}^n \hat{\varepsilon}_i \cdot \hat{\varepsilon}_{i-1}}{\sum_{i=1}^n \hat{\varepsilon}_i^2}$$

La relation liant la statistique DW et l'estimateur $\hat{\rho}$ permet de se faire une idée de la valeur vraisemblable du coefficient d'autocorrélation ρ à partir de la valeur prise par la statistique DW .

En effet :

- Si $\rho = 0$ et donc s'il n'y a pas d'autocorrélation des perturbations, on s'attend à ce que $\hat{\rho}$ soit proche de 0. La statistique DW devrait alors prendre une valeur proche de 2,

² Durbin, J. et G.S. Watson « *Testing for serial correlation in least-squared regression* », *Biometrika*, 1951.

- En cas d'autocorrélation positive maximale ($\rho = 1$), on s'attend à ce que $\hat{\rho}$ soit proche de 1. Dans ce cas, la statistique DW devrait prendre une valeur proche de 0,
- Inversement, en cas d'autocorrélation négative maximale ($\rho = -1$), on s'attend à ce que $\hat{\rho}$ soit proche de -1 . Dans ce cas, la statistique DW devrait prendre une valeur proche de 4.

Pour effectuer le test à partir de la statistique DW , il faut déterminer une valeur critique de cette statistique et donc connaître la distribution de probabilité de DW . Or, cette distribution de probabilité dépend des valeurs des observations des variables explicatives figurant dans la matrice X . Ainsi, il est impossible de construire des tables statistiques permettant de connaître, dans tous les cas de figure, les valeurs critiques de DW , pour différents seuils de significativité.

Durbin et Watson ont montré que, **sous l'hypothèse nulle d'absence d'autocorrélation, il est possible d'encadrer la statistique DW par les valeurs prises par deux autres variables aléatoires**, notées en général d_L et d_U (avec $d_L < d_U$) dont les lois ne dépendent que des paramètres connus à savoir n , le nombre d'observations et ($k' = k - 1$), le nombre de variables explicatives du modèle.

La procédure à adopter pour mener le test de Durbin-Watson est la suivante :

1. Appliquer les moindres carrés ordinaires au modèle initial et en déduire la valeur prise par la statistique DW . Cette information fait partie des sorties automatiques de la plupart des logiciels d'économétrie ;
2. Pour une valeur α du risque de première espèce, chercher dans une table de Durbin-Watson les valeurs critiques $d_{L,\alpha}^*$ et $d_{U,\alpha}^*$. Ci-dessous un extrait d'une table de Durbin-Watson pour un risque de première espèce égal à 5% :

TABLE 7b Durbin-Watson Statistic: 5 Percent Significance Points of dL and dU										
n	$k' = 1$		$k' = 2$		$k' = 3$		$k' = 4$		$k' = 5$	
	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
6	0.610	1.400	—	—	—	—	—	—	—	—
7	0.700	1.356	0.467	1.896	—	—	—	—	—	—
8	0.763	1.332	0.559	1.777	0.368	2.287	—	—	—	—
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588	—	—
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822
11	0.927	1.324	0.758	1.604	0.595	1.928	0.444	2.283	0.316	2.645
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.355	1.771
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772
85	1.624	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774
90	1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.571	1.780
150	1.720	1.746	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802
200	1.758	1.778	1.748	1.789	1.738	1.799	1.728	1.810	1.718	1.820

Figure 4 : extrait d'une table de Durbin-Watson pour un risque de première espèce égal à 5%

3. Décision :

- Si $DW < d_{L,\alpha}^*$: on rejette H_0 et on considère que $\rho > 0$,
- Si $DW > 4 - d_{L,\alpha}^*$: on rejette H_0 et on considère que $\rho < 0$,
- Si $d_{U,\alpha}^* < DW < 4 - d_{U,\alpha}^*$: on ne rejette pas H_0 et on considère que $\rho = 0$,
- Si $d_{L,\alpha}^* < DW < d_{U,\alpha}^*$ ou si $4 - d_{U,\alpha}^* < DW < 4 - d_{L,\alpha}^*$: le test ne permet pas de conclure.

Astuce

Pour éviter les erreurs, il est pratique de reproduire le schéma suivant en y reportant les valeurs critiques correspondant au problème :

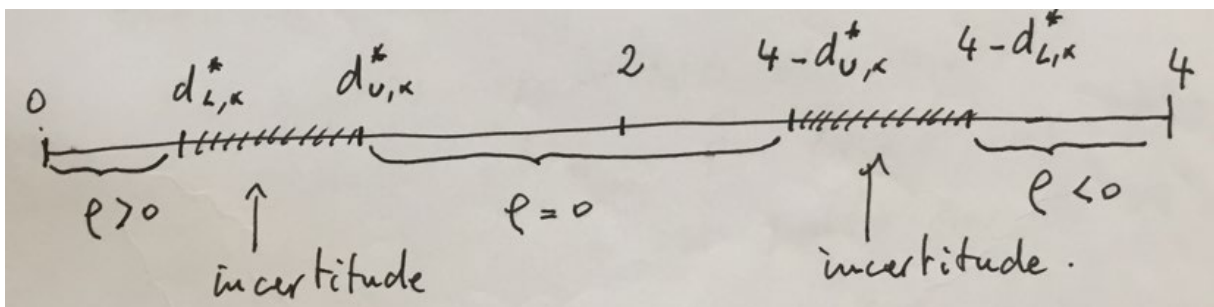


Figure 5 : schéma des valeurs critiques correspondant au problème

Remarques à propos du test de Durbin-Watson

- Ce test ne permet d'identifier qu'une autocorrélation d'ordre 1 ;
- La variable dépendante ne doit pas figurer dans la liste des variables explicatives en tant que variable retardée ;
- Le modèle doit impérativement comporter un terme constant ;
- Pour les modèles en coupe transversale, les observations doivent être ordonnées en fonction de la variable à expliquer.

Le test de Breusch³ – Godfrey⁴

Ce test est fondé sur un test de Fisher de nullité de coefficients ou sur un test du multiplicateur de Lagrange (test LM). Le double intérêt de ce test est qu'il **permet de tester une autocorrélation d'un ordre supérieur à 1** et qu'il reste valide en présence de la variable dépendante décalée en tant que variable explicative du modèle (deux des limitations énoncées pour le test de Durbin-Watson).

L'idée générale de ce test réside dans la **recherche d'une relation significative entre le résidu estimé de l'estimation et ce même résidu décalé.**

Une autocorrélation des erreurs d'ordre p s'écrit :

$$\varepsilon_i = \rho_1 \varepsilon_{i-1} + \rho_2 \varepsilon_{i-2} + \dots + \rho_p \varepsilon_{i-p} + u_i$$

où u_i est une perturbation gaussienne d'espérance nulle et de variance constante.

Le modèle linéaire multiple à erreurs auto corrélées d'ordre p s'écrit donc, pour l'observation i :

$$y_i = \beta_0 + \beta_1 \cdot x_{1,i} + \beta_2 \cdot x_{2,i} + \dots + \beta_{k-1} \cdot x_{k-1,i} + \rho_1 \varepsilon_{i-1} + \rho_2 \varepsilon_{i-2} + \dots + \rho_p \varepsilon_{i-p} + u_i$$

³ Breusch, T. « *Testing for autocorrelation in dynamic linear models* », Australian Economic Papers, 17, 1978.

⁴ Godfrey, L. « *Testing against general autoregressive and moving average errors models when the regressors include lagged dependant variables* », Econometrica, 46, 1978.

La procédure à adopter pour mener le test de Breusch-Godfrey est la suivante :

1. Estimation par moindres carrés ordinaires du modèle initial et calcul des résidus estimés $\hat{\varepsilon}_i$ ($i = 1, \dots, n$). Puisque les erreurs sont inconnues, le test porte sur les résidus estimés,
2. Estimation par moindres carrés ordinaires de l'équation intermédiaire :

$$\hat{\varepsilon}_i = \beta_0 + \beta_1 \cdot x_{1,i} + \beta_2 \cdot x_{2,i} + \dots + \beta_{k-1} \cdot x_{k-1,i} + \rho_1 \hat{\varepsilon}_{i-1} + \rho_2 \hat{\varepsilon}_{i-2} + \dots + \rho_p \hat{\varepsilon}_{i-p} + u_i$$

Soit T le nombre d'observations disponibles pour estimer les paramètres de l'équation intermédiaire (Attention ! Chaque décalage entraîne la perte d'une observation) et soit R_{ε}^2 le coefficient de détermination associé à cette dernière estimation.

3. Test d'hypothèse sur l'équation intermédiaire : l'hypothèse nulle d'absence d'autocorrélation des erreurs s'écrit :

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_p = 0$$

Si l'on rejette cette hypothèse nulle, alors il existe un risque d'autocorrélation des erreurs à l'ordre p .

Pour mener ce test on a deux possibilités :

- Soit effectuer un test de Fisher classique de nullité des coefficients ρ_j ,
- Soit recourir à la statistique $LM = T \cdot R_{\varepsilon}^2$ qui, sous l'hypothèse nulle, est distribuée selon une variable du Khi-deux à p degrés de liberté. Si $T \cdot R_{\varepsilon}^2$ est supérieure à la valeur critique lue dans la table au seuil α , on rejette l'hypothèse d'indépendance des erreurs.

Conséquences de l'autocorrélation des erreurs sur la méthode des MCO

Les conséquences sont identiques à celles décrites pour la situation d'hétéroscédasticité des perturbations, à savoir :

- Les estimateurs des MCO définis par $\hat{\beta} = (X'X)^{-1}X'y$ restent sans biais en présence d'autocorrélation des erreurs ;

- L'expression $\text{Var}_{(n,n)} \varepsilon = \sigma^2 \cdot I_n$ ne permet plus d'obtenir de bons estimateurs des variances des coefficients estimés. Les résultats des tests qui les utilisent ne sont donc plus valides ;
- Les estimateurs des Moindres Carrés Ordinaires ne sont plus des estimateurs BLUE.

Estimation d'un modèle en présence d'autocorrélation des perturbations

Souvent, la présence d'autocorrélation des résidus d'une estimation révèle plus un problème au niveau de la spécification retenue que d'une simple autocorrélation des erreurs. La présence d'autocorrélation des résidus peut être due :

- À l'oubli d'une variable explicative importante lors de l'estimation ;
- Au choix d'une mauvaise forme fonctionnelle pour l'équation estimée.

Il en résulte que le premier réflexe à avoir en présence d'autocorrélation des résidus estimés est de **revoir la spécification retenue**, en envisageant d'introduire d'autres variables explicatives ou de choisir d'autres formes fonctionnelles lors de l'estimation.

Si le problème persiste, c'est-à-dire si l'autocorrélation semble être un problème authentique, il est nécessaire d'adopter la méthode décrite ci-après.

Méthode à adopter quand l'autocorrélation semble être un problème authentique

Pour faciliter la présentation, on va envisager ici le cas d'une autocorrélation des perturbations d'ordre 1 dans le cadre d'un modèle de régression simple.

Soit le modèle à estimer suivant :

$$y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i \quad i = 1, \dots, n$$

où les perturbations sont autocorrélées selon un processus $AR(1)$ c'est-à-dire :

$$\varepsilon_i = \rho \cdot \varepsilon_{i-1} + u_i \quad \text{avec } |\rho| < 1 ; u_i \rightarrow LN(0 ; \sigma^2) ; Cov(\varepsilon_i, u_i) = 0$$

Supposons que le coefficient d'autocorrélation ρ soit connu de l'économètre. Dans ce cas, une simple transformation du modèle initial permet d'obtenir des perturbations sphériques.

Écrivons le modèle pour les observations i et $i - 1$:

$$y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i \quad (A) \quad i = 1, \dots, n$$

$$y_{i-1} = \beta_0 + \beta_1 \cdot x_{i-1} + \varepsilon_{i-1} \quad (B) \quad i = 2, \dots, n$$

Retranchons ρ fois la relation (B) à la relation (A) :

$$y_i - \rho \cdot y_{i-1} = \beta_0 \cdot (1 - \rho) + \beta_1 \cdot (x_i - \rho \cdot x_{i-1}) + \varepsilon_i - \rho \cdot \varepsilon_{i-1}$$

$$\Leftrightarrow y_i^* = \beta_0 \cdot (1 - \rho) + \beta_1 \cdot x_i^* + u_i \quad (C) \quad i = 2, \dots, n$$

où $y_i^* = y_i - \rho \cdot y_{i-1}$ et $x_i^* = x_i - \rho \cdot x_{i-1}$

Par construction, les perturbations du modèle (C) sont sphériques. On dira que l'on a écrit le modèle initial en **quasi-différences**. L'estimation par Moindres Carrés Ordinaires du coefficient β_1 à partir du modèle transformé (C) s'interprète directement comme étant le coefficient estimé du modèle initial.

Cependant, dans la pratique, le coefficient d'autocorrélation ρ n'est pas connu. Il convient donc de l'estimer.

Plusieurs méthodes existent pour estimer ρ .

Estimation de ρ par régression des résidus estimés du modèle initial $\hat{\varepsilon}_i$ sur leurs valeurs retardées $\hat{\varepsilon}_{i-1}$

$$\hat{\rho} = \frac{Cov(\hat{\varepsilon}_i, \hat{\varepsilon}_{i-1})}{Var \hat{\varepsilon}_i} = \frac{\sum_{i=2}^n \hat{\varepsilon}_i \cdot \hat{\varepsilon}_{i-1}}{\sum_{i=1}^n \hat{\varepsilon}_i^2}$$

Estimation de ρ à partir de la statistique de Durbin-Watson, DW

$$\hat{\rho} = 1 - \frac{DW}{2}$$

Méthode de Durbin⁵

Le modèle en quasi-différences s'écrivait :

$$y_i - \rho \cdot y_{i-1} = \beta_0 \cdot (1 - \rho) + \beta_1 \cdot (x_i - \rho \cdot x_{i-1}) + \varepsilon_i - \rho \cdot \varepsilon_{i-1}$$

⁵ Durbin, J. « *The fitting of time series models* », Review of the International Statistical Institute, 28, 1960.

$$\Leftrightarrow y_i = \rho \cdot y_{i-1} + \beta_0 \cdot (1 - \rho) + \beta_1 \cdot x_i - \beta_1 \rho \cdot x_{i-1} + u_i$$

Il n'est pas nécessaire de connaître ρ pour estimer ce dernier modèle qui peut être réécrit sous la forme :

$$y_i = \rho \cdot y_{i-1} + \gamma_0 + \beta_1 \cdot x_i + \gamma_1 \cdot x_{i-1} + u_i$$

Puisqu'à présent la méthode des moindres carrés ordinaires est applicable (perturbations sphériques), l'estimation du coefficient de y_{i-1} fournit une estimation de ρ .

Méthode de Cochrane-Orcutt⁶

Cette méthode se présente sous la forme d'une procédure itérative, chaque étape produisant une meilleure estimation de ρ que la précédente. Ici encore, la méthode sera présentée dans le cadre d'un modèle de régression simple, mais elle se généralise sans problème au cas d'un modèle ayant plusieurs variables explicatives.

Soit le modèle linéaire simple sous la forme :

$$y_i = a + b \cdot x_i + \varepsilon_i \quad i = 1, \dots, n$$

La procédure à adopter pour appliquer la méthode de Cochrane-Orcutt est la suivante :

1. Estimer par moindres carrés ordinaires le modèle initial et en déduire les résidus estimés $\hat{\varepsilon}_i$, où $\hat{\varepsilon}_i = y_i - \hat{y}_i = y_i - \hat{a} - \hat{b}x_i$.
2. Utiliser ces résidus estimés pour estimer l'équation autorégressive :

$$\hat{\varepsilon}_i = \rho \cdot \hat{\varepsilon}_{i-1} + u_i$$

En déduire la valeur estimée du coefficient d'autocorrélation :

$$\hat{\rho} = \frac{Cov(\hat{\varepsilon}_i, \hat{\varepsilon}_{i-1})}{Var \hat{\varepsilon}_i} = \frac{\sum_{i=2}^n \hat{\varepsilon}_i \cdot \hat{\varepsilon}_{i-1}}{\sum_{i=1}^n \hat{\varepsilon}_i^2}$$

3. Utiliser cette estimation de ρ pour spécifier le modèle en quasi-différences de la façon suivante :

$$y_i - \hat{\rho} \cdot y_{i-1} = a \cdot (1 - \hat{\rho}) + b \cdot (x_i - \hat{\rho} \cdot x_{i-1}) + u_i$$

⁶ Cochrane, D. et G. Orcutt « *Application of least squares regression to relationships containing autocorrelated error terms* », Journal of the American Statistical Association, 44, 1969.

$$\Leftrightarrow y_i^* = a^* + b^* \cdot x_i^* + u_i$$

Estimer ce modèle par MCO et en déduire les estimations des paramètres du modèle initial :

$$\hat{a}_1 = \frac{\hat{a}^*}{1 - \hat{\rho}} \quad ; \quad \hat{b}_1 = \hat{b}^*$$

4. Substituer les nouveaux paramètres estimés dans le modèle initial et en déduire la valeur des nouveaux résidus estimés $\hat{\varepsilon}_i$ où :

$$\hat{\varepsilon}_i = y_i - \hat{a}_1 - \hat{b}_1 x_i$$

5. À partir de ces nouveaux résidus estimés, reprendre la procédure à l'étape 2. Il faut donc estimer à nouveau l'équation autorégressive :

$$\hat{\varepsilon}_i = \rho \cdot \hat{\varepsilon}_{i-1} + u_i$$

et en déduire une nouvelle estimation du coefficient d'autocorrélation ρ , et ainsi de suite.

Cette procédure itérative peut être poursuivie autant de fois que l'on veut. En général, la règle est d'arrêter la procédure lorsque la nouvelle estimation de ρ diffère de la précédente de moins de 10^{-2} .

Méthode du balayage de Hildreth-Lu⁷

Dans cette méthode, on va donner au coefficient d'autocorrélation ρ des valeurs successives. On retiendra celle qui est compatible avec le critère sous-jacent aux MCO.

Le schéma employé est le suivant :

1. À partir de la statistique de Durbin-Watson, on détermine le type d'autocorrélation susceptible d'être présent (positive ou négative – DW proche de 0 ou proche de 4).
2. Si, par exemple, on soupçonne une autocorrélation positive, on va donner à ρ une suite de valeurs dans l'intervalle $[0,1]$, en se fixant un certain pas a-priori. Par exemple, pour un pas fixé à 10^{-1} , on donnera à ρ les valeurs suivantes :

$$\rho = \{0; 0,1; 0,2; \dots; 0,9; 1\}$$

⁷ Hildreth, C. et J. Lu « Demand relations with autocorrelated disturbances », Technical Bulletin N° 276, Michigan State University Agricultural Experiment Station, 1960.

Pour chacune de ces valeurs de ρ , on estime le modèle en quasi-différences :

$$y_i^* = a^* + b^* \cdot x_i^* + u_i$$

où $y_i^* = y_i - \rho \cdot y_{i-1}$ et $x_i^* = x_i - \rho \cdot x_{i-1}$. La procédure consiste à retenir la valeur de ρ correspondant à l'équation ayant la plus faible somme des carrés des résidus estimés ($\sum_i \hat{u}_i^2$).

3. Une fois cette valeur de ρ déterminée, il est possible d'affiner la valeur estimée de ρ en réemployant la même procédure sur un intervalle restreint et avec un pas plus petit (10^{-2} par exemple).

En conclusion, quelle que soit la méthode employée pour estimer ρ , on calcule les estimations des coefficients du modèle en quasi-différences en remplaçant ρ par son estimation précédemment obtenue.

Références

Comment citer ce cours ?

Introduction à l'économétrie, Olivier Baron, AUNEGe (<http://aunega.fr>), CC – BY NC ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Cette œuvre est mise à disposition dans le respect de la législation française protégeant le droit d'auteur, selon les termes du contrat de licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). En cas de conflit entre la législation française et les termes de ce contrat de licence, la clause non conforme à la législation française est réputée non écrite. Si la clause constitue un élément déterminant de l'engagement des parties ou de l'une d'elles, sa nullité emporte celle du contrat de licence tout entier.

Table des illustrations

Figures

Figure 1 : autocorrélation positive des résidus estimés	4
Figure 2 : : autocorrélation négative des résidus estimés.....	5
Figure 3 : graphique des résidus estimés.....	6
Figure 4 : extrait d'une table de Durbin-Watson pour un risque de première espèce égal à 5% 10	
Figure 5 : schéma des valeurs critiques correspondant au problème	10