Tableaux jointement échangemables dissociés
Application économétrique

Yanis Zeghal (M1)

Encadré par

Laurent Davezies et Raphaël Lafrogne-Joussier

3 mai 2023

Ce rapport propose une méthode économétrique pour des modèles à données jointement échangeables et dissociées. Particulièrement adaptés à l'étude du commerce bilatéral ou de réseaux, ces modèles imposent des conditions d'indépendance plus faibles que les conditions $i.i.d$ et différentes de l'indépendance entre clusters. L.Davezies, X.d'Haultfœuille et Y.Guyonvarch [DDG22] proposent en 2022 des théorèmes asymptotiques et de nouvelles méthodes économétriques spécifiques aux modèles avec données jointement échangeables et dissociées. Ce rapport introduit les notions nécessaires, présente des outils économétriques et propose une démonstration de tous les résultats utilisés. Une implémentation en R est présentée dans un second temps. Cet outil a été utilisé pour valider les résultats de l'étude de Frankel et Rose [FR98] par des étudiants de l'ENSAE lors d'un projet de statistiques appliquées.

Table des matières

1 Préliminaires et motivation du modèle

$(Y_{i_1...,i_d})_{i_1...i_d \in \mathbb{N}^*}$ est un tableau à d indices dans $\mathbb{N}$. C'est une variable aléatoire. On considère que les indices sont tous distincts. La notation rigoureuse est alors : $(Y_{i_1,...,i_d})_{(i_1,...,i_d)\in\mathbb{N}^{*d},\forall k\neq k':i_k\neq i_{k'}}$. Par souci de simplicité, nous utiliserons la première notation. On définit les deux propriétés suivantes :

Définition 1

Echangeabilité jointe : On dit que $(Y_{i_1...,i_d})_{i_1...i_d \in \mathbb{N}^*}$ est jointement échangeable si pour toute permutation π de $\mathbb{N}^*$:

$$(Y_{\pi(i_1),\pi(i_2),...,\pi(i_d)})_{(i_1,...,i_d)\in\mathbb{N}^{*d},\forall k\neq k':i_k\neq i_{k'}} \overset{loi}{=} (Y_{i_1,...,i_d})_{(i_1,...,i_d)\in\mathbb{N}^{*d},\forall k\neq k':i_k\neq i_{k'}} \tag{1}$$

Définition 2

Dissociation : On dit que $(Y_{i_1,...,i_d})_{i_1...i_d \in \mathbb{N}^*}$ est dissocié si pour tout $n > 0$:

$$(Y_{i_1,...,i_d})_{(i_1,...,i_d)\in[\![1,n]\!]^d,\forall k\neq k':i_k\neq i_{k'}} \perp\!\!\!\perp (Y_{i_1,...,i_d})_{(i_1,...,i_d)\in[\![n+1,\infty[\![^d,\forall k\neq k':i_k\neq i_{k'}} \tag{2}$$

On dit que $(Y_{i_1...,i_d})_{i_1...i_d \in \mathbb{N}^*}$ est un tableau jointement échangeable et dissocié s'il remplit ces deux conditions. On remarque alors trois propriétés importantes, qui justifient notre choix d'utilisation de ce modèle :

Propriété 0

Tous les éléments de $(Y_{i_1,...,i_d})_{i_1...i_d \in \mathbb{N}^*}$ suivent la même loi, qui est la loi de $Y_{1...,d}$.
Preuve : Pour une d-liste $(i_1, ..., i_d)$ donnée , prenons une permutation π telle que :
$\pi : \mathbb{N}^* \to \mathbb{N}^*$
$i_1 \mapsto 1$
$i_2 \mapsto 2$
...
$i_d \mapsto d$
Puis appliquons l'équation (1). Les lois de Y avant et après permutations des indices étant les mêmes, les lois marginales de Y sont aussi inchangées. Il en résulte que :

$$Y_{i_1,...,i_d} \overset{loi}{=} Y_{\pi(i_1),\pi(i_2),...,\pi(i_d)} \overset{loi}{=} Y_{1,...,d} \tag{3}$$

Propriété 1

Prenons deux d-listes d'indices $(i_1, ..., i_d)$ et $(j_1, ..., j_d) \in \mathbb{N}^d$ n'ayant pas d'indice en commun : $\forall t, t' \in [\![1, d]\!] : i_t \neq j_{t'}$; Alors $Y_{i_1,...,i_d} \perp\!\!\!\perp Y_{j_1,...,j_d}$.
Preuve : Pour s'en rendre compte, on peut construire une permutation π telle que $\pi(\{i_1, ..., i_d\}) = [\![1, d]\!]$ On a donc $\forall t \in [\![1, d]\!], \pi(j_t) \geq d + 1$
Par l'équation (1), le tableau obtenu après permutation des indices a la même loi que la tableau obtenu avant permutation, puis en appliquant l'équation (2) avec $n = d$, on obtient que $Y_{i_1,...,i_d} \perp\!\!\!\perp Y_{j_1,...,j_d}$, car ils suivent les lois marginales de variables indépendantes.

Propriété 2

Ce raisonnement sera utilisé à plusieurs reprises dans les démonstrations. Pour $Y_{i,j}$ un tableau jointement échangeable et dissocié à 2 indices et $a, b, c \in \mathbb{N}^*$ distincts :
Les quadruplets de forme $(Y_{a,b}, Y_{b,a}, Y_{a,c}, Y_{c,a})$ ont la même loi que $(Y_{1,2}, Y_{2,1}, Y_{1,3}, Y_{3,1})$.
preuve : Prendre une permutation π telle que :
$\pi : \mathbb{N}^* \to \mathbb{N}^*$
$a \mapsto 1$
$b \mapsto 2$
$c \mapsto 3$

Propriété 3

Choisissons $p < d$ indices (sans restriction de généralité, prenons les p premiers indices) et considérons les couples ayant ces p indices en commun, et n'ayant pas d'indice commun parmi les $d - p$ indices restants :
$(Y_{a_1,a_2,...,a_p,i_{p+1},...,i_d}\ ,\ Y_{a_1,a_2,...,a_p,j_{p+1},...,j_d})$ tel que $\{i_{p+1}, ..., i_d\} \bigcap \{j_{p+1}, ..., j_d\} = \varnothing$. En considérant de nouveau que ce couple suit une loi marginale du tableau $(Y_{i_1...,i_d})_{i_1...i_d \in \mathbb{N}^*}$, et par l'équation (1), ce couple a la même loi que $(Y_{1,2,...,p,...d}\ ,\ Y_{1,2,...,p,d+1...2*d-p+1})$ et que tous les couples de la même forme (ces p indices identiques, et le reste des indices disjoints). On montre ceci de la même manière que précédemment, en choisissant une permutation qui envoie les p premiers indices sur $[\![1, p]\!]$, les $\{i_{p+1}...i_d\}$ sur $\{p + 1...d\}$ et $\{j_{p+1}...j_d\}$ sur $\{d + 1...2 * d - p + 1\}$.

Motivation

Ces propriétés motivent l'utilisation du modèle. Il est courant d'être confronté à un jeu de données avec 2 indices. Dans l'étude de Frankel & Rose ([FR98]), les paires de pays forment un tableau à 2 indices i et j, qui sont distincts (il n'y a pas d'observation pour le commerce bilatéral d'un pays à lui-même). Un modèle de tableau jointement échangeable et dissocié est adéquat pour les raisons suivantes :
Les observations sont indépendantes si elles n'ont pas d'indice commun (propriété 1). Ainsi, une observation pour le couple (*France-Allemagne*) n'est pas supposée indépendante d'une observation pour le couple (*France-Belgique*) car elles ont l'indice *France* en commun, mais est indépendante d'une observation pour le couple (*Italie - Royaume-Uni*), car elles n'ont pas d'indices commun. Dans ce cas, ce modèle définit les hypothèses de ce qui est indépendant ou non d'une manière plus réaliste que permet le clustering sur le premier ou second indice.
En appliquant la propriété 3 aux tableaux à 2 indices, les couples d'observation ayant un indice en commun sont identiquement corrélées pour tous les pays. En termes économiques, cela signifie que les pays font des choix de partenaires économiques suivant les mêmes mécanismes, que l'on peut justifier en supposant qu'ils maximisent des fonctions d'utilité similaires. Cela reste une hypothèse.

Lorsqu'on est confrontés à des données sur des flux entre utilisateurs de réseaux, il peut être adéquat d'utiliser les hypothèses d'observations jointement échangeables et dissociées. Pour un groupe de n personnes, les échanges entre i et j ne sont pas indépendants des échanges entre i et k mais sont indépendants des échanges entre k et l, et tous les flux suivent la même loi.

Définition de l'outil

Pour les MCO On suppose que $(Y_{i,j}, X_{i,j})_{i,j}$ est un tableau jointement échangeable et dissocié tel que $Y_{i,j} = X'_{i,j}\beta_0 + \varepsilon_{i,j}$ avec $K = \dim(X) = \dim(\beta_0)$, $J = \mathbb{E}(X_{1,2}X'_{1,2})$ est inversible, $\mathbb{E}(Y_{1,2}^4 + ||X_{1,2}||^4) < \infty$ et $\mathbb{E}(X_{i,j}\varepsilon_{i,j}) = 0$. On note $\widehat{\beta}$ l'estimateur des MCO et $\widehat{\varepsilon}_{i,j} = Y_{i,j} - X'_{i,j}\widehat{\beta}$. On montre en section 2.5 les formules suivantes :

$$\sqrt{n}\left[\widehat{J}^{-1}\widehat{H}\widehat{J}^{-1}\right]^{-1/2}\left(\widehat{\beta} - \beta_0\right) \xrightarrow{d} \mathcal{N}\left(0, Id_K\right) \tag{4}$$

Avec :

$$\widehat{J} = \frac{1}{n(n-1)}\sum_{i,j} X_{i,j}X'_{i,j} \tag{5}$$

$$\widehat{H} = \frac{1}{n(n-1)^2}\sum_{i}\left(\sum_{j\neq i} X_{i,j}\widehat{\varepsilon}_{i,j} + X_{j,i}\widehat{\varepsilon}_{j,i}\right)\left(\sum_{j\neq i} X_{i,j}\widehat{\varepsilon}_{i,j} + X_{j,i}\widehat{\varepsilon}_{j,i}\right)' \tag{6}$$

Pour les DMC On suppose maintenant que $(Y_{i,j}, X_{i,j}, Z_{i,j})$ est un tableau jointement échangeable et dissocié tel que $Y_{i,j} = X'_{i,j}\beta_0 + \varepsilon_{i,j}$ avec $K = \dim(X) = \dim(\beta_0)$, $A = \mathbb{E}(Z_{1,2}Z'_{1,2})$ est inversible, $B = \mathbb{E}(Z_{1,2}X'_{1,2})$ est de plein rang colonne, $\mathbb{E}(Y_{1,2}^4 + ||Z_{1,2}||^4 + ||X_{1,2}||^2) < \infty$ et $\mathbb{E}(Z_{i,j}\varepsilon_{i,j}) = 0$. On note $\widehat{\beta}$ l'estimateur des doubles moindres carrés et $\widehat{\varepsilon}_{i,j} = Y_{i,j} - X'_{i,j}\widehat{\beta}$. On montre en section 2.6 les formules suivantes :

$$\sqrt{n}\left[\left(\widehat{B}'\widehat{A}^{-1}\widehat{B}\right)^{-1}\widehat{B}\widehat{A}^{-1}\widehat{H}\widehat{A}^{-1}\widehat{B}\left(\widehat{B}'\widehat{A}^{-1}\widehat{B}\right)^{-1}\right]^{-1/2}\left(\widehat{\beta} - \beta_0\right) \xrightarrow{d} \mathcal{N}\left(0, Id_K\right) \tag{7}$$

Avec :

$$\widehat{A} = \frac{1}{n(n-1)}\sum_{i,j} Z_{i,j}Z'_{i,j} \tag{8}$$

$$\widehat{B} = \frac{1}{n(n-1)}\sum_{i,j} Z_{i,j}X'_{i,j} \tag{9}$$

$$\widehat{H} = \frac{1}{n(n-1)^2}\sum_{i}\left(\sum_{j\neq i} Z_{i,j}\widehat{\varepsilon}_{i,j} + Z_{j,i}\widehat{\varepsilon}_{j,i}\right)\left(\sum_{j\neq i} Z_{i,j}\widehat{\varepsilon}_{i,j} + Z_{j,i}\widehat{\varepsilon}_{j,i}\right)' \tag{10}$$

2 Démonstrations

2.1 Représentation d'Aldous-Hoover-Kallenberg

La démonstration des théorèmes énoncés repose sur la représentation de Aldous-Hoover-Kallenberg pour les tableaux jointement échangeables dissociés à 2 indices. Soit des variables aléatoires $(U_i)_i$ et $(U_{i,j})_{i,j}$ telles que :

les $(U_i)_i$ sont $i.i.d$

les $(U_{i,j})_{i,j}$ sont $i.i.d$

$(U_i)_i \perp\!\!\!\perp (U_{i,j})_{i,j}$

les U_i et $U_{i,j}$ sont uniformes sur $[0, 1]$

Soit f une fonction mesurable de $[0, 1]^3$ dans $\mathbb{R}^k$. Le tableau $f(U_i, U_j, U_{\{i,j\}})_{i,j}$ est un tableau jointement échangeable et dissocié, car il remplit les conditions des définitions 1 et 2.

Echangeabilité jointe :

Soit π une permutation de $\mathbb{N}^*$. En appliquant π aux indices de $(U_i, U_j, U_{i,j})$, on obtient :
$(U_{\pi(i)}, U_{\pi(j)}, U_{\pi(i),\pi(j)})_{i,j} \overset{loi}{=} (U_i, U_j, U_{i,j})_{i,j}$
Puis
$\left(f(U_{\pi(i)}, U_{\pi(j)}, U_{\pi(i),\pi(j)})\right)_{i,j} \overset{loi}{=} (f(U_i, U_j, U_{i,j}))_{i,j}$

Dissociation :

Soit $n \in \mathbb{N}*$.
$(U_i, U_j, U_{i,j})_{i,j \leq n} \perp\!\!\!\perp (U_i, U_j, U_{i,j})_{i,j>n}$
Donc :
$(f(U_i, U_j, U_{i,j}))_{i,j \leq n} \perp\!\!\!\perp (f(U_i, U_j, U_{i,j}))_{i,j>n}$

[Kal05] permet d'admettre la réciproque [1] :
Si $(Y_{i,j})_{i,j \in \mathbb{N}^{*2}, i \neq j}$ est un tableau jointement échangeable et dissocié alors il existe une fonction f mesurable et des $(U_i)_i$ et $(U_{i,j})_{i,j}$ tels que :

$$Y_{i,j} = f(U_i, U_j, U_{\{i,j\}}).$$

1. Ceci est le seul résultat admis dans ce rapport

Cette écriture est la représentation d'Aldous-Hoover-Kallenbeg [2].
On note $g(u) = \mathbb{E}\left[f(u, U_j, U_{\{i,j\}}) + f(U_j, u, U_{\{i,j\}})\right] = \mathbb{E}\left(Y_{i,j} + Y_{j,i} | U_i = u\right)$.
Remarquons que la définition de g ne dépend pas des indices i et j. En effet, tous les quadruplets de forme $(Y_{a,b}, Y_{b,a}, Y_{a,c}, Y_{c,a})$ ont bien la même loi que $(Y_{1,2}, Y_{2,1}, Y_{1,3}, Y_{3,1})$, en vertu de la propriété 2. Il en va donc de même pour les $(Y_{a,b}, Y_{b,a})$.

Le résultat suivant sera utile par la suite :

$$\mathrm{Cov}\left(Y_{12} + Y_{21}, Y_{13} + Y_{31}\right) = \mathbb{V}\left(\mathbb{E}\left(Y_{12} + Y_{21} | U_1\right)\right) = \mathbb{V}(g(U_1)) \tag{11}$$

On le démontre en utilisant la décomposition de la covariance (pour des vecteurs aléatoires) :

$$\mathrm{Cov}\left(Y_{12} + Y_{21}, Y_{13} + Y_{31}\right) =$$

$$\mathrm{Cov}\left(\mathbb{E}\left(Y_{12} + Y_{21} | U_1\right), E\left(Y_{13} + Y_{31} | U_1\right)\right) + \mathbb{E}\left(\underbrace{\mathrm{Cov}\left(Y_{12} + Y_{21}, Y_{13} + Y_{31} | \, U_1\right)}_{=0}\right)$$

En remarquant que le terme dans l'espérance est nul car $(Y_{12} + Y_{21}, Y_{13} + Y_{31}) = (f(U_1, U_2, U_{\{1,2\}}) + f(U_2, U_1, U_{\{2,1\}}), f(U_1, U_2, U_{\{1,2\}}) + f(U_2, U_1, U_{\{2,1\}}))$. *Conditionnellement à* U_1, ces deux variables aléatoires sont seulement fonction de U_2 et $U_{1,2}$ d'une part et U_3 et $U_{1,3}$ d'autre part. Elles sont donc indépendantes et leur covariance *conditionnelle à* U_1 est nulle.
Pour le premier terme, on remarque que $\mathbb{E}(Y_{12} + Y_{21} | U_1) = E(Y_{13} + Y_{31} | U_1)$, puisque ce sont les espérances d'une même fonction de U_2 et U_3 à U_1 fixé. On peut donc transformer la covariance en variance et obtenir l'équation ci-dessus.

2.2 Théorème de la Limite Centrale

Dans cette partie, on montre un théorème central-limite pour les tableaux à 2 indices.
Pour $\mathbb{E}(Y_{i,j}^2) < \infty$:

$$\sqrt{n}\left(\frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} Y_{i,j} - \mathbb{E}\left(Y_{1,2}\right)\right) \xrightarrow{d} \mathcal{N}(0, \mathbb{V}(g(U)))$$

2. En général, ni la fonction f, ni les $(U_i)_i$ et $(U_{i,j})_{i,j}$ sont uniques. La fonction f peut être très irrégulière. Toutes les manières de construire une variable uniforme à partir de deux autres variables uniformes peuvent servir d'exemple.

Premièrement, si $\mathbb{E}(Y_{1,2}) = 0$ on montre l'égalité suivante :

$$\begin{aligned} &\mathbb{E}\left[\left|\left|\frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n} Y_{i,j} - \frac{1}{n}\sum_{i=1}^{n} g\left(U_i\right)\right|\right|^2\right] \\ =&\frac{1}{n(n-1)}\left[\mathbb{E}\left(\frac{||Y_{12}+Y_{21}||^2}{2}\right) - \mathbb{E}\left(||g(U_1)||^2\right)\right] \end{aligned} \tag{12}$$

Pour ce faire, on applique le théorème de pythagore à des variables aléatoires astucieusement regroupées. Premièrement, les termes $Y_{i,j} + Y_{j,i} - g(U_i) - g(U_j)$ sont orthogonaux deux à deux.

— S'ils n'ont pas d'indice commun, ce sont des fonctions de variables déjà indépendantes, et de plus :
  $\mathbb{E}(Y_{i,j} + Y_{j,i} - g(U_i) - g(U_j)) = \underbrace{\mathbb{E}(Y_{i,j} + Y_{j,i})}_{=0} - \underbrace{\mathbb{E}(g(U_i) - g(U_j))}_{=0}$

— S'ils ont un indice a commun : En notant$< | >$ le produit scalaire sur l'espace des vecteurs aléatoires, le terme $\mathbb{E}((Y_{a,i}+Y_{i,a})g(U_a))$ s'écrit $< Y_{a,i}+Y_{i,a}|g(U_a) >$ puis par construction, $g(U_a)$ est le projeté orthogonal de $Y_{a,i} + Y_{i,a}$ sur $\sigma(U_a)$, d'où :

$$\begin{aligned} < Y_{a,i} + Y_{i,a}|g(U_a) >&= < g(U_a)|g(U_a) > + \underbrace{< Y_{a,i} + Y_{i,a} - g(U_a)|g(U_a) >}_{=0} \\ &= < g(U_a)|g(U_a) > \\ &=\mathbb{E}(g(U_a)^2) \end{aligned}$$

Ce qui permet d'établir, (parce qu'ils ont des indices i et j différents) :

$$\begin{aligned} &\mathbb{E}((Y_{a,i} + Y_{i,a} - g(U_i) - g(U_a))(Y_{a,j} + Y_{j,a} - g(U_j) - g(U_a))) \\ =&\underbrace{\mathbb{E}((Y_{a,i} + Y_{i,a})(Y_{a,j} + Y_{j,a}))}_{=\mathbb{E}(g(U_a)^2)\text{par eq11}} + \mathbb{E}(g(U_a)^2) - 2\,\underbrace{\mathbb{E}((Y_{a,i} + Y_{i,a})g(U_a))}_{=\mathbb{E}(g(U_a)^2)} \\ =&0 \end{aligned}$$

On peut ensuite regrouper les termes comme suit :

$$
\begin{aligned}
&\mathbb{E}\left[\left|\left|\frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n} Y_{i,j} - \frac{1}{n}\sum_{i=1}^{n} g\left(U_i\right)\right|\right|^2\right] \\
=&\mathbb{E}\left[\left|\left|\frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n} \left(Y_{i,j} - g\left(U_i\right)\right)\right|\right|^2\right] \\
=&\mathbb{E}\left[\left|\left|\frac{1}{n(n-1)}\Big(\sum_{i\leq n}\sum_{j<i}\left(Y_{i,j} - g\left(U_i\right)\right) + \sum_{i\leq n}\sum_{j>i}\left(Y_{i,j} - g\left(U_i\right)\right)\Big)\right|\right|^2\right]
\end{aligned}
$$

Puis en renommant les indices i et j dans le deuxième terme, on obtient :

$$
\begin{aligned}
=&\mathbb{E}\left[\left|\left|\frac{1}{n(n-1)}\Big(\sum_{i\leq n}\sum_{j<i}\left(Y_{i,j} - g\left(U_i\right)\right) + \sum_{i\leq n}\sum_{j<i}\left(Y_{j,i} - g\left(U_j\right)\right)\Big)\right|\right|^2\right] \\
=&\mathbb{E}\left[\left|\left|\frac{1}{n(n-1)}\Big(\sum_{i\leq n}\sum_{j<i}\left(Y_{i,j} + Y_{j,i} - g\left(U_i\right) - g\left(U_i\right)\right)\right|\right|^2\right]
\end{aligned}
$$

Tous les termes étant orthogonaux 2 à 2, on peut appliquer le théorème de pythagore (plusieurs fois) et écrire :

$$
\begin{aligned}
=&\frac{1}{(n(n-1))^2}\sum_{i=2}^{n}\sum_{j<i}\mathbb{E}\left[||(Y_{i,j} + Y_{j,i} - g\left(U_i\right) - g\left(U_i\right)||^2\right] \\
=&\frac{1}{2n(n-1)}\mathbb{E}\left[||Y_{1,2} + Y_{2,1} - g\left(U_1\right) - g\left(U_2\right)||^2\right] \\
=&\frac{1}{n(n-1)}\left[\mathbb{E}\left(\frac{||Y_{1,2} + Y_{2,1}||^2}{2}\right) - \mathbb{E}\left[||g\left(U_1\right)||^2\right]\right]
\end{aligned}
$$

L'équation (12) donne une convergence L^2 (donc en probabilité et en loi) en $O(\frac{1}{n^2})$. Dans le cas où $\mathbb{E}(Y_{1,2}) \neq 0$, on peut remarquer que $(Y_{i,j} - \mathbb{E}(Y_{1,2}))_{i,j}$ est toujours un tableau jointement échangeable dissocié. De plus, sa représentation d'Aldous Hoover Kallenberg est obtenue par translation de la fonction des variables centrées :

$Y_{i,j} - \mathbb{E}(Y_{1,2}) = f(U_i, U_j, U_{i,j})$
et
$\mathbb{E}(Y_{i,j} + Y_{j,i}|U_i) = \mathbb{E}(Y_{i,j} - \mathbb{E}(Y_{1,2}) + Y_{j,i} - \mathbb{E}(Y_{1,2})|U_i) + 2\mathbb{E}(Y_{1,2}) = g(U_i) + 2\mathbb{E}(Y_{1,2})$.

On gardera la notation g pour les variables centrées et $g + 2\mathbb{E}(Y_{1,2})$ pour les variables quelconques.

$$\frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} (Y_{i,j} - \mathbb{E}(Y_{1,2})) = \left(\frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} Y_{i,j} - \mathbb{E}(Y_{1,2}) - \frac{1}{n} \sum_{i=1}^{n} \underbrace{g(U_i)}_{g\text{centrées}} \right) + \frac{1}{n} \sum_{i=1}^{n} g(U_i)$$

Le premier terme est exactement celui de l'équation (12). Le second est une somme de variables i.i.d. Puis en multipliant par $\sqrt{n}$, on obtient :

$$\frac{\sqrt{n}}{n(n-1)} \sum_{1 \leq i \neq j \leq n} (Y_{i,j} - \mathbb{E}(Y_{1,2})) = \underbrace{\sqrt{n} \left(\frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} Y_{i,j} - \mathbb{E}(Y_{1,2}) - \frac{1}{n} \sum_{i=1}^{n} g(U_i) \right)}_{O(n^{\frac{-3}{2}})}$$
$$+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} g(U_i)$$

Le premier terme tend vers 0 en probabilité, puisque c'est un $O(n^{\frac{-3}{2}})$ en norme 2. Sous l'hypothèse $\mathbb{E}(||g(U_1)||^2) < \infty$, le second tend en loi vers $\mathcal{N}(0, \mathbb{V}(g(U_1)))$. Le théorème de Slutsky permet alors de conclure :

$$\sqrt{n} \left(\frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} Y_{i,j} - \mathbb{E}\left(Y_{1,2}\right) \right) \xrightarrow{d} \mathcal{N}(0, \mathbb{V}(g(U))) \tag{13}$$

2.3 Loi des grands nombres pour $d \geq 2$

On s'intéresse maintenant à un tableau dissocié jointement échangeable de dimension d : $(Y_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \text{distincts}}$. On montre dans cette section que si $\mathbb{E}\left(||Y_{1,\ldots,d}||\right) < \infty$ alors

$$\frac{(n-d)!}{n!} \sum_{i_1,\ldots,i_d \leq n} Y_{i_1,\ldots,i_d} \xrightarrow{P} \mathbb{E}(Y_{1,\ldots,d})$$

Sans restreindre la généralité, nous prouverons nos résultats pour des variables aléatoires dans $\mathbb{R}$, puisque la convergence en probabilité de toutes les composantes d'une

suite de vecteurs aléatoires est équivalente à la convergence en probabilité de la suite elle-même.
On montre ici la convergence :

$$\frac{(n-d)!}{n!} \sum_{i_1,\dots,i_d \le n} Y_{i_1,\dots,i_d} \xrightarrow{L1} \mathbb{E}\left(Y_{1,\dots,d}\right) \tag{14}$$

Supposons dans un premier temps que $\exists M \in \mathbb{R} \quad tq \quad |Y_{1,2,\dots,d}| < M \quad p.s$, et que $\mathbb{E}(Y_{1,2,\dots,d}) = 0$. On montre alors que $\frac{(n-d)!}{n!} \sum_{i_1,\dots,i_d \le n} Y_{i_1,\dots,i_d}$ converge dans L^2.

$$\begin{aligned}
\left\| \frac{(n-d)!}{n!} \sum_{i_1,\dots,i_d \le n} Y_{i_1,\dots,i_d} \right\|_{L^2}^2 &= \left(\frac{(n-d)!}{n!}\right)^2 \mathbb{E}\left(\sum_{\substack{i_1,\dots,i_d \le n \\ j_1,\dots,j_d \le n}} Y_{i_1,\dots,i_d} Y_{j_1,\dots,j_d} \right) \\
&= \left(\frac{(n-d)!}{n!}\right)^2 \sum_{\substack{i_1,\dots,i_d \le n \\ j_1,\dots,j_d \le n}} \mathbb{E}(Y_{i_1,\dots,i_d} Y_{j_1,\dots,j_d})
\end{aligned}$$

La propriété 1 permet d'écrire $\mathbb{E}(Y_{i_1,\dots,i_d} Y_{j_1,\dots,j_d}) = 0$ dès lors que $(i_1, \dots, i_d)$ et $(j_1, \dots, j_d)$ n'on pas d'indices commun. On peut ensuite majorer les autres termes : $\mathbb{E}(|Y_{i_1,\dots,i_d} Y_{j_1,\dots,j_d}|) \le M^2$. Le nombre de termes nuls dans la somme est donné par le nombre de $2d$ - listes d'éléments pris dans $\{1, \dots, n\}$ soit $\frac{n!}{(n-2d)!}$. Le nobre de termes non nuls est donc :

$$\underbrace{(\frac{n!}{(n-d)!})^2}_{\text{nb total de termes}} - \underbrace{\frac{n!}{(n-2d)!}}_{\text{nb termes nuls}}$$

D'où :

$$\left(\frac{(n-d)!}{n!}\right)^2 \sum_{\substack{i_1,\dots,i_d \le n \\ j_1,\dots,j_d \le n}} \mathbb{E}(|Y_{i_1,\dots,i_d} Y_{j_1,\dots,j_d}|) \le \left(1 - \underbrace{\frac{(n-d)!^2}{n!(n-2d)!}}_{\to 1} \right) M \xrightarrow[n\to\infty]{} 0$$

Cette convergence L^2 implique la convergence L^1 par l'inégalité de Hölder. Dans le cas où $Y_{1,\dots,d}$ n'est pas majorée presque sûrement, on montre que :
$\forall \eta > 0 \quad \exists M > 0 \quad \text{tq} \quad \mathbb{E}(|Y_{1,\dots,d}| \, \mathbb{1}_{|Y_{1,\dots,d}|>M}) < \eta$.
Pour cela, on peut remarquer que la suite de variables aléatoires $(|Y_{1,\dots,d}| \mathbb{1}_{|Y_{1,\dots,d}|>M})_{M\in\mathbb{N}}$ est majorée par $|Y_{1,\dots,d}| \in L^1$ et converge $p.s$ vers 0. Le théorème de convergence dominée (ou le théorème de convegrence monotone) donne $E\left(|Y_{1,\dots,d} \mathbb{1}_{|Y_{1,\dots,d}|>M}|\right) \xrightarrow[M\to\infty]{} 0$

On peut ensuite décomposer :

$$\begin{aligned}Y_{i_1,\dots,i_d} =& Y_{i_1,\dots,i_d}\mathbb{1}\{|Y_{i_1,\dots,i_d}| \le M\} - \mathbb{E}\left[Y_{1,\dots,d}\mathbb{1}\{|Y_{1,\dots,d}| \le M\}\right] \\ &+ Y_{i_1,\dots,i_d}\mathbb{1}\{|Y_{i_1,\dots,i_d}| > M\} - \mathbb{E}\left[Y_{1,\dots,d}\mathbb{1}\{|Y_{1,\dots,d}| > M\}\right],\end{aligned}$$

Pour la deuxième ligne, l'inégalité triangulaire permet d'écrire :

$$\mathbb{E}\left(|Y_{i_1,\dots,i_d}\mathbb{1}\{|Y_{i_1,\dots,i_d}| > M\} - \mathbb{E}\left[Y_{1,\dots,d}\mathbb{1}\{|Y_{1,\dots,d}| > M\}\right]|\right) \le 2\mathbb{E}\left[|Y_{1,\dots,d}|\,\mathbb{1}\{|Y_{1,\dots,d}| > M\}\right]$$

puis

$$\left(\frac{(n-d)!}{n!}\right)\sum_{i_1,\dots,i_d}\mathbb{E}\left(|Y_{i_1,\dots,i_d}\mathbb{1}\{|Y_{i_1,\dots,i_d}| > M\} - \mathbb{E}\left[Y_{1,\dots,d}\mathbb{1}\{|Y_{1,\dots,d}| > M\}\right]|\right) \le \frac{\eta}{2}$$

pour M assez grand et pour tout $n \ge d$

Le premier terme $Y_{i_1,\dots,i_d}\mathbb{1}\{|Y_{i_1,\dots,i_d}| \le M\} - \mathbb{E}\left[Y_{1,\dots,d}\mathbb{1}\{|Y_{1,\dots,d}| \le M\}\right]$ est le terme général d'un tableau jointement échangeable dissocié. La convergence (14) appliquée aux variables majorées en valeurs absolues et d'espérance nulle permet d'écrire :

$$\mathbb{E}\left[\left|\left(\frac{(n-d)!}{n!}\right)\sum_{i_1,\dots,i_d}Y_{i_1,\dots,i_d}\mathbb{1}\{|Y_{i_1,\dots,i_d}| \le M\} - \mathbb{E}\left[Y_{1,\dots,d}\mathbb{1}\{|Y_{1,\dots,d}| \le M\}\right]\right|\right] \le \frac{\eta}{2}$$

pour n assez grand, avec M choisi ci-dessus

Ces deux inégalités étant vraies pour tout η, on a :

$$\begin{aligned}&\limsup_{n\to\infty}\mathbb{E}\left[\left|\left(\frac{(n-d)!}{n!}\right)\sum_{i_1,\dots,i_d}Y_{i_1,\dots,i_d}\right|\right] = 0 \\ \Rightarrow\quad & \frac{(n-d)!}{n!}\sum_{i_1,\dots,i_d}Y_{i_1,\dots,i_d} \xrightarrow{L^1} 0 \\ \Rightarrow\quad & \frac{(n-d)!}{n!}\sum_{i_1,\dots,i_d}Y_{i_1,\dots,i_d} \xrightarrow{P} 0\end{aligned}$$

Pour un $Y_{1,\dots,d}$ non centré, on peut appliquer cette convergence à $Y_{i_1,\dots,i_d} - \mathbb{E}(Y_{1,\dots,d})$ (qui est le terme général d'un tableau jointement échangeable dissocié) pour obtenir :

$$\frac{(n-d)!}{n!}\sum_{i_1,\dots,i_d}Y_{i_1,\dots,i_d} \xrightarrow{P} \mathbb{E}(Y_{1,\dots,n}) \tag{15}$$

On remarque que cette convergence en probabilité généralise celle obtenue par l'égalité (12) aux variables aléatoires de L^1 (et non plus L^2) et à d et non plus 2 indices.

2.4 Inférence pour les tableaux à 2 indices

Dans cette partie, nous construisons un intervalle de confiance pour la moyenne du terme général d'un tableau jointement échangeable dissocié.
Soit $(Y_{i,j})_{i\neq j}$ un tableau jointement échangeable dissocié à 2 indices, avec $Y_{1,2} \in \mathbb{R}^d$ et $\mathbb{E}(||Y_{1,2}||^2) < \infty$.
On montre que $((Y_{i,j}+Y_{j,i})(Y_{i,k}+Y_{k,i})')_{i,j,k \text{ distincts}}$ est un tableau jointement échangeable dissocié à 3 indices.
On utilise la représentation d'Aldous Hoover Kallenberg de $(Y_{i,j})_{i\neq j}$.

Echangeabilité jointe :
: $((Y_{i,j}+Y_{j,i})(Y_{i,k}+Y_{k,i})')_{i,j,k}$ est fonction de $(U_i, U_j, U_k, U_{i,j}, U_{i,k})_{i,j,k}$.
En appliquant une permutation π à ses indices, la loi suivante reste inchangée :
$(U_{\pi(i)}, U_{\pi(j)}, U_{\pi(k)}, U_{\pi(i),\pi(j)}, U_{\pi(i),\pi(k)})_{i,j,k} \overset{loi}{=} (U_i, U_j, U_k, U_{i,j}, U_{i,k})_{i,j,k}$
parce que leur définition ne change pas (ce sont des lois uniformes $i.i.d$ sur $[0,1]$, avec la même indexation).
Donc :
$((Y_{i,j}+Y_{j,i})(Y_{i,k}+Y_{k,i})')_{i,j,k} \overset{loi}{=} \left((Y_{\pi(i),\pi(j)}+Y_{\pi(j),\pi(i)})(Y_{\pi(i),\pi(k)}+Y_{\pi(k),\pi(i)})'\right)_{i,j,k}$

Dissociation : $\forall n \in \mathbb{N}$, comme $(Y_{i,j})_{i,j}$ est dissocié : $((Y_{i,j}, Y_{j,i}, Y_{i,k}, Y_{k,i}))_{i,j,k\leq n} \perp\!\!\!\perp ((Y_{i,j}, Y_{j,i}, Y_{i,k}, Y_{k,i}))_{i,j,k>n}$.
Puis
$((Y_{i,j}+Y_{j,i})(Y_{i,k}+Y_{k,i})')_{i,j,k\leq n} \perp\!\!\!\perp ((Y_{i,j}+Y_{j,i})(Y_{i,k}+Y_{k,i})')_{i,j,k>n}$

Dans un premier temps, supposons $\mathbb{E}(Y_{1,2}) = 0$. On montre le résultat suivant :

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}\sum_{j\neq i} Y_{i,j}+Y_{j,i}\right)\left(\frac{1}{n-1}\sum_{j\neq i} Y_{i,j}+Y_{j,i}\right)' \xrightarrow{P} \mathbb{V}(g(U_1)) \tag{16}$$

La somme peut être décomposée :

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}\sum_{j\neq i} Y_{i,j}+Y_{j,i}\right)\left(\frac{1}{n-1}\sum_{j\neq i} Y_{i,j}+Y_{j,i}\right)' \\
=&\frac{1}{n(n-1)^2}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{k\neq i,j}\left(Y_{i,j}+Y_{j,i}\right)\left(Y_{i,k}+Y_{k,i}\right)' && \left(\xrightarrow{P} \mathbb{V}(g(U_1))\right) \\
+&\frac{1}{n(n-1)^2}\sum_{i=1}^{n}\sum_{j\neq i}\left(Y_{i,j}+Y_{j,i}\right)\left(Y_{i,j}+Y_{j,i}\right)' && \left(\xrightarrow{P} 0\right)
\end{aligned}$$

Le premier terme converge vers $\mathbb{E}\left(((Y_{1,2}+Y_{2,1})(Y_{1,3}+Y_{3,1})')\right)$ par l'équation (15), qui s'écrit aussi $\mathbb{V}(g(U_1))$ par l'équation (11).
Le second terme converge vers 0 en remarquant que $\left((Y_{i,j}+Y_{j,i})\left(Y_{i,j}+Y_{j,i}\right)'\right)_{i,j}$ est aussi un tableau jointement échangeable et dissocié, de terme général intégrable en norme et en appliquant l'équation (15). Ceci conclut la preuve pour des variables centrées.
Pour des variables non centrées, on applique cette convergence à $Y_{i,j}-\mathbb{E}(Y_{1,2})$:

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\mathbb{E}(Y_{1,2})\right)\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\mathbb{E}(Y_{1,2})\right)'\xrightarrow{P}\mathbb{V}(g(U_1))$$

En remarquant que les fonctions g définies pour les variables $Y_{i,j}$ et $Y_{i,j}-\mathbb{E}(Y_{1,2})$ ont la même variance. Pour construire un estimateur de la variance dans le cas non centré, on note $\overline{Y}=\frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n}Y_{i,j}$. Puis :

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\mathbb{E}(Y)\right)\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\mathbb{E}(Y)\right)'\\
=&\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\overline{Y}-(2\mathbb{E}(Y)-2\overline{Y})\right)\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\overline{Y}-(2\mathbb{E}(Y)-2\overline{Y})\right)'\\
=&\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\overline{Y}\right)\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\overline{Y}\right)'\\
&-\left(2\mathbb{E}(Y)-2\overline{Y}\right)\underbrace{\left(\frac{1}{n(n-1)}\sum_{i=1}^{n}(Y_{i,j}+Y_{j,i})-2\overline{Y}\right)'}_{=0} && =0\\
&-\underbrace{\left(\frac{1}{n(n-1)}\sum_{i=1}^{n}(Y_{i,j}+Y_{j,i})-2\overline{Y}\right)}_{=0}\left(2\mathbb{E}(Y)-2\overline{Y}\right)' && =0\\
&+\left(2\mathbb{E}(Y)-2\overline{Y}\right)\left(2\mathbb{E}(Y)-2\overline{Y}\right)' && \xrightarrow{P}0\textit{(théorème de continuité)}
\end{aligned}$$

D'où

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\overline{Y}\right)\left(\frac{1}{n-1}\sum_{j\neq i}Y_{i,j}+Y_{j,i}-2\overline{Y}\right)'\xrightarrow{P}\mathbb{V}(g(U_1))\quad(17)$$

Cette convergence donne un estimateur consistant de $V(g(U_1))$, et permet de procéder à des tests asymptotiques de niveau α pour $\mathbb{E}(Y_{1,2})$, en reprenant la convergence (13). On note $\widehat{V_a}$ l'estimateur de $V(g(U_1))$ ci-dessus.

$$(13) \Rightarrow \widehat{V_a}^{\frac{-1}{2}} \left(\frac{\sqrt{n}}{n(n-1)} \sum_{1 \leq i \neq j \leq n} Y_{i,j} - \mathbb{E}\left(Y_{1,2}\right) \right) \xrightarrow{d} \mathcal{N}(0, I_d)$$

Un intervalle de confiance de niveau α pour l'estimateur de $\mathbb{E}(Y1,2)$ peut être construit pour tester une hypothèse d'égalité sur un coefficient :
$H_0 \quad : \quad \mathbb{E}(Y_{1,2}) = b$
$H_1 \quad : \quad \mathbb{E}(Y_{1,2}) \neq b$

Sous H_0 $\quad \lim_{n\to\infty} \mathbb{P}_{H_0}\left(\overline{Y} \in \left[b - \frac{q_{1-\alpha/2}}{\sqrt{n}\widehat{V}_{a,j,j}^{-1/2}}, b + \frac{q_{1-\alpha/2}}{\sqrt{n}\widehat{V}_{a,j,j}^{-1/2}}\right]\right) = 1-\alpha$

Sous H_1 $\quad \lim_{n\to\infty} \mathbb{P}_{H_0}\left(\overline{Y} \in \left[b - \frac{q_{1-\alpha/2}}{\sqrt{n}\widehat{V}_{a,j,j}^{-1/2}}, b + \frac{q_{1-\alpha/2}}{\sqrt{n}\widehat{V}_{a,j,j}^{-1/2}}\right]\right) = 0$

2.5 Inférence pour les Moindres Carrés Ordinaires

Dans le cas d'un tableau jointement échangeable dissocié $(Y_{i,j}, X_{i,j})_{i,j}$ tel que $Y_{i,j} = X'_{i,j}\beta_0 + \varepsilon_{i,j}$ avec $K = \dim(X) = \dim(\beta_0)$, on note $J = \mathbb{E}(X_{1,2}X'_{1,2})$ (que l'on suppose inversible) et on suppose $\mathbb{E}(Y_{1,2}^4 + ||X_{1,2}||^4) < \infty$ et $\mathbb{E}(X_{i,j}\varepsilon_{i,j}) = 0$. On note $\widehat{\beta}$ l'estimateur des MCO et $\widehat{\varepsilon}_{i,j} = Y_{i,j} - X'_{i,j}\widehat{\beta}$.

Les lois marginales $(X_{i,j})_{i,j}$, $(Y_{i,j})_{i,j}$ et les fonctions mesurables de ces variables (y compris $(\varepsilon_{i,j})_{i,j}$) sont des tableaux jointement échangeables dissociés. On définit :

$$\widehat{J} = \frac{1}{n(n-1)} \sum_{i,j} X_{i,j}X'_{i,j}$$

La convergence (15) donne :

$$\widehat{J} = \frac{1}{n(n-1)} \sum_{i,j} X_{i,j}X'_{i,j} \xrightarrow{P} \mathbb{E}(X_{1,2}X'_{1,2}) = J \tag{18}$$

On définit :

$$\widetilde{H} = \frac{1}{n(n-1)^2} \sum_{i} \left(\sum_{j\neq i} X_{i,j}\varepsilon_{i,j} + X_{j,i}\varepsilon_{j,i} \right) \left(\sum_{j\neq i} X_{i,j}\varepsilon_{i,j} + X_{j,i}\varepsilon_{j,i} \right)$$

La convergence (16) appliquée à $(X_{i,j}\varepsilon_{i,j})_{i,j}$ donne :

$$\frac{1}{n(n-1)^2}\sum_i\left(\sum_{j\neq i}X_{i,j}\varepsilon_{i,j}+X_{j,i}\varepsilon_{j,i}\right)\left(\sum_{j\neq i}X_{i,j}\varepsilon_{i,j}+X_{j,i}\varepsilon_{j,i}\right)\xrightarrow{P}\mathbb{E}\left((X_{12}\varepsilon_{12}+X_{21}\varepsilon_{21})\left(X_{13}\varepsilon_{13}+X_{31}\varepsilon_{31}\right)'\right)$$

On note H cette limite, et donc $\widetilde{H}\xrightarrow{P}H$.

L'estimateur des moindres carrés ordinaires vérifie :

$$\widehat{\beta}-\beta_0=\widehat{J}^{-1}\frac{1}{n(n-1)}\sum_{i,j}X_{i,j}\varepsilon_{i,j}$$

Comme $(X_{i,j}\varepsilon_{i,j})_{i,j}$ est jointement échangeable dissocié, la convergence (16) et le théorème de Slutsky donnent directement $\widehat{\beta}-\beta_0\xrightarrow{P}0$.
De plus, la convergence (13) donne :

$$\frac{\sqrt{n}}{n(n-1)}\sum_{i,j}X_{i,j}\varepsilon_{i,j}\xrightarrow{d}\mathcal{N}\left(0,\underbrace{\mathbb{E}\left((X_{12}\varepsilon_{12}+X_{21}\varepsilon_{21})\left(X_{13}\varepsilon_{13}+X_{31}\varepsilon_{31}\right)'\right)}_{=H}\right)$$

D'où :

$$\sqrt{n}\left(\widehat{\beta}-\beta_0\right)\xrightarrow{d}\mathcal{N}\left(0,J^{-1}HJ^{-1}\right)\tag{19}$$

Afin de pouvoir estimer la variance asymptotique, on souhaite utiliser $\widetilde{H}$ en substituant des $\widehat{\varepsilon}_{i,j}$ aux $\varepsilon_{i,j}$.
On note $\widehat{H}$ la statistique :

$$\widehat{H}=\frac{1}{n(n-1)^2}\sum_i\left(\sum_{j\neq i}X_{i,j}\widehat{\varepsilon}_{i,j}+X_{j,i}\widehat{\varepsilon}_{j,i}\right)\left(\sum_{j\neq i}X_{i,j}\widehat{\varepsilon}_{i,j}+X_{j,i}\widehat{\varepsilon}_{j,i}\right)$$

En écrivant

$$\widehat{\varepsilon}_{i,j}-\varepsilon_{i,j}=(Y_{i,j}-X'_{i,j}\widehat{\beta})-(Y_{i,j}-X'_{i,j}\beta_0)=X'_{i,j}(\beta_0-\widehat{\beta})$$

On peut décomposer $\widehat{H}$ comme suit :

$$\begin{aligned}
\widehat{H} =& \frac{1}{n(n-1)^2}\sum_i\left(\sum_{j\neq i}X_{i,j}\varepsilon_{i,j}+X_{j,i}\varepsilon_{j,i}\right)\left(\sum_{j\neq i}X_{i,j}\varepsilon_{i,j}+X_{j,i}\varepsilon_{j,i}\right)\\
&+\frac{1}{n(n-1)^2}\sum_i\left(\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right)\left(\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right)'\\
&+\frac{1}{n(n-1)^2}\sum_i\left(\sum_{j\neq i}X_{i,j}\varepsilon_{i,j}+X_{i,j}\varepsilon_{i,j}\right)\left(\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right)'\\
&+\frac{1}{n(n-1)^2}\sum_i\left(\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right)\left(\sum_{j\neq i}X_{i,j}\varepsilon_{i,j}+X_{i,j}\varepsilon_{i,j}\right)'
\end{aligned} \tag{$*$}$$

Le premier terme est $\widetilde{H}$. On montre que les trois autres termes convergent en probabilité vers 0. La convergence du deuxième terme se montre grâce à la norme de Frobenius définie par $||A||_F^2=\sum_{i,j}a_{i,j}^2$ sur l'espace des matrices réelles (de dimension quelconque). On utilise la propriété suivante :

$$\forall A\in\mathcal{M}_{1,K}:\quad ||AA'||_F=\sqrt{\sum_{i,j}(a_i^2a_j^2)}=\sum_i a_i^2=||A||_2^2$$

Comme $\forall i:\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\in\mathcal{M}_{1,K}$, on obtient :

$$\begin{aligned}
&\frac{1}{n(n-1)^2}\left|\left|\sum_i\left(\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right)\left(\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right)'\right|\right|_F\\
\leq&\frac{1}{n(n-1)^2}\sum_i\left|\left|\left(\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right)\left(\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right)'\right|\right|_F\\
\leq&\frac{1}{n(n-1)^2}\sum_i\left|\left|\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right|\right|_2^2
\end{aligned}$$

En considérant $\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)$ comme la matrice d'une application linéaire,
$\left|\left|\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\left(\widehat{\beta}-\beta_0\right)\right|\right|_2\leq\left|\left|\sum_{j\neq i}\left(X_{i,j}X_{i,j}'+X_{j,i}X_{j,i}'\right)\right|\right|_{\mathbb{R}^K\to\mathbb{R}^K}\left|\left|\left(\widehat{\beta}-\beta_0\right)\right|\right|_2$.

La norme $||\cdot||_{\mathbb{R}^K\to\mathbb{R}^K}$ étant équivalente à toute autre norme matricielle (on est dans un espace de dimension $K^2<\infty$), on peut choisir à nouveau la norme de Frobenius, et
:

$$\exists C \in \mathbb{R}+ : \quad \begin{aligned} \left|\left|\sum_{j\neq i}\left(X_{i,j}X'_{i,j}+X_{j,i}X'_{j,i}\right)\right|\right|_{\mathbb{R}^K\to\mathbb{R}^K} \leq & C\left|\left|\sum_{j\neq i}\left(X_{i,j}X'_{i,j}+X_{j,i}X'_{j,i}\right)\right|\right|_F \\ \leq & C\sum_{j\neq i}\left(\left|\left|X_{i,j}X'_{i,j}\right|\right|_F+\left|\left|X_{j,i}X'_{j,i}\right|\right|_F\right) \\ = & C\sum_{j\neq i}\left(||X_{i,j}||_2^2+||X_{j,i}||_2^2\right) \end{aligned}$$

D'où :

$$\left|\left|\sum_{j\neq i}\left(X_{i,j}X'_{i,j}+X_{j,i}X'_{j,i}\right)\left(\widehat{\beta}-\beta_0\right)\right|\right|_2^2 \leq C^2\left[\sum_{j\neq i}\left(||X_{i,j}||_2^2+||X_{i,j}||_2^2\right)\right]^2\left|\left|\left(\widehat{\beta}-\beta_0\right)\right|\right|_2^2$$

Donc :

$$\begin{aligned} & \frac{1}{n(n-1)^2}\sum_i\left|\left|\sum_{j\neq i}\left(X_{i,j}X'_{i,j}+X_{j,i}X'_{j,i}\right)\left(\widehat{\beta}-\beta_0\right)\right|\right|_2^2 \\ \leq & \underbrace{\frac{C^2}{n(n-1)^2}\sum_i\left[\sum_{j\neq i}\left(||X_{i,j}||_2^2+||X_{j,i}||_2^2\right)\right]^2}_{\text{Majoré en espérance car } \mathbb{E}(||X_{1,2}||^4)<\infty}\underbrace{\left|\left|\left(\widehat{\beta}-\beta_0\right)\right|\right|_2^2}_{\xrightarrow{P}0} \end{aligned}$$

Les deux derniers termes de la décomposition $(*)$ étant transposés l'un de l'autre, on montre leur convergence vers 0 pour le deuxième seulement.
Pour $(A,B)\in\mathcal{M}_{1,n}\times\mathcal{M}_{1,m}$:

$$||AB'||_F=\sum_{i=1}^n\sum_{j=1}^m a_i^2b_j^2=\sum_{i=1}^n a_i^2\sum_{j=1}^m b_j^2=||A||_2||B||_2$$

La norme du dernier terme est :

$$\frac{1}{n(n-1)^2}\left|\left|\sum_i\left(\sum_{j\neq i}\left(X_{i,j}X'_{i,j}+X_{j,i}X'_{j,i}\right)\left(\widehat{\beta}-\beta_0\right)\right)\left(\sum_{j\neq i}X_{i,j}\varepsilon_{i,j}+X_{i,j}\varepsilon_{i,j}\right)'\right|\right|_F$$

$$\leq\frac{1}{n(n-1)^2}\sum_i\left|\left|\left(\sum_{j\neq i}\left(X_{i,j}X'_{i,j}+X_{j,i}X'_{j,i}\right)\left(\widehat{\beta}-\beta_0\right)\right)\left(\sum_{j\neq i}X_{i,j}\varepsilon_{i,j}+X_{i,j}\varepsilon_{i,j}\right)'\right|\right|_F$$

$$\leq\frac{1}{n(n-1)^2}\sum_i\underbrace{\left|\left|\left(\sum_{j\neq i}\left(X_{i,j}X'_{i,j}+X_{j,i}X'_{j,i}\right)\left(\widehat{\beta}-\beta_0\right)\right)\right|\right|_2}_{\leq C\left[\sum_{j\neq i}\left(||X_{i,j}||_2^2+||X_{i,j}||_2^2\right)\right]\left|\left|\left(\widehat{\beta}-\beta_0\right)\right|\right|_2}\left|\left|\left(\sum_{j\neq i}X_{i,j}\varepsilon_{i,j}+X_{i,j}\varepsilon_{i,j}\right)\right|\right|_2$$

$$\leq\underbrace{\frac{C}{n(n-1)^2}\sum_i\left(\left(\sum_{j\neq i}||X_{i,j}||_2^2+||X_{i,j}||_2^2\right)\left(\sum_{j\neq i}||X_{i,j}\varepsilon_{i,j}||_2+||X_{i,j}\varepsilon_{i,j}||_2\right)\right)}_{\text{Majoré en espérance}}\underbrace{\left|\left|\left(\widehat{\beta}-\beta_0\right)\right|\right|_2}_{\xrightarrow{P}0}$$

Comme $X_{1,2}$ et $Y_{1,2}$ sont dans L^4, ε est aussi dans L^4, et la majoration en espérance du terme de gauche est possible grâce à l'inégalité de Cauchy-Schwarz et parce qu'aucun terme de la somme ne dépasse le degré 4.
Cette dernière convergence permet de conclure :

$$\widehat{H}-\widetilde{H}\xrightarrow{P}0\text{ et donc }\widehat{H}\xrightarrow{P}H \tag{20}$$

Les convergences (19), (18) et (20), le théorème de continuité et le théorème de Slutsky donnent la convergence 4 :

$$\sqrt{n}\left[\widehat{J}^{-1}\widehat{H}\widehat{J}^{-1}\right]^{-1/2}\left(\widehat{\beta}-\beta_0\right)\xrightarrow{d}\mathcal{N}\left(0,Id_K\right)$$

2.6 Inférence pour les Doubles Moindres Carrés

Dans le cas d'un tableau jointement échangeable dissocié $(Y_{i,j},X_{i,j},Z_{i,j})$ tel que $Y_{i,j}=X'_{i,j}\beta_0+\varepsilon_{i,j}$ avec $K=\dim(X)=\dim(\beta_0)$, on note $A=\mathbb{E}(Z_{1,2}Z'_{1,2})$ (que l'on suppose inversible), et $B=\mathbb{E}(Z_{1,2}X'_{1,2})$ que l'on suppose de plein rang colonne. Les variables $\mathbb{E}(Y_{1,2}^4+||Z_{1,2}||^4+||X_{1,2}||^4)<\infty$ et $\mathbb{E}(Z_{i,j}\varepsilon_{i,j})=0$. On note $\widehat{\beta}$ l'estimateur des doubles moindres carrés et $\widehat{\varepsilon}_{i,j}=Y_{i,j}-X'_{i,j}\widehat{\beta}$.
Les étapes de cette démonstration se font de la même manière que précédemment :
On montre grâce à la convergence (15) que :

$$\widehat{A}=\frac{1}{n(n-1)}\sum_{i,j}Z_{i,j}Z'_{i,j}\xrightarrow{P}A$$

Comme $(Z_{i,j}X'_{i,j})_{i,j}$ est aussi un tableau jointement échangeable dissocié, on montre en utilisant la convergence (15) une nouvelle fois que :

$$\widehat{B} = \frac{1}{n(n-1)} \sum_{i,j} Z_{i,j}X'_{i,j} \xrightarrow{P} B$$

La convergence (16) appliquée à $(Z_{i,j}\varepsilon_{i,j})_{i,j}$ donne :

$$\widetilde{H} = \frac{1}{n(n-1)^2} \sum_i \left(\sum_{j \neq i} Z_{i,j}\varepsilon_{i,j} + Z_{j,i}\varepsilon_{j,i} \right) \left(\sum_{j \neq i} Z_{i,j}\varepsilon_{i,j} + Z_{j,i}\varepsilon_{j,i} \right)' \xrightarrow{P} H$$

avec

$$H = \mathbb{E}\left((Z_{12}\varepsilon_{12} + Z_{21}\varepsilon_{21}) (Z_{13}\varepsilon_{13} + Z_{31}\varepsilon_{31})' \right)$$

La définition de l'estimateur des doubles moindres carrés et la convergence (13) permettent d'obtenir :

$$\sqrt{n}\left(\widehat{\beta} - \beta_0\right) \xrightarrow{d} \mathcal{N}\left(0, (B'A^{-1}B)^{-1}B'A^{-1}HA^{-1}B(B'A^{-1}B)^{-1}\right).$$

Par les mêmes arguments que dans la partie sur les moindres carrés ordinaires (section 2.5), mais avec des calculs plus lourds, on peut montrer que :

$$\widehat{H} = \frac{1}{n(n-1)^2} \sum_i \left(\sum_{j \neq i} Z_{i,j}\widehat{\varepsilon}_{i,j} + Z_{j,i}\widehat{\varepsilon}_{j,i} \right) \left(\sum_{j \neq i} Z_{i,j}\widehat{\varepsilon}_{i,j} + Z_{j,i}\widehat{\varepsilon}_{j,i} \right)' \xrightarrow{P} H$$

Le théorème de continuité et le théorème de Slutsky et les équations ci-dessus donnent la convergence (7) :

$$\sqrt{n}\left[\left(\widehat{B}'\widehat{A}^{-1}\widehat{B}\right)^{-1} \widehat{B}\widehat{A}^{-1}\widehat{H}\widehat{A}^{-1}\widehat{B} \left(\widehat{B}'\widehat{A}^{-1}\widehat{B}\right)^{-1} \right]^{-1/2} \left(\widehat{\beta} - \beta_0\right) \xrightarrow{d} \mathcal{N}\left(0, Id_K\right).$$

□

2.7 Limites et extensions

Une limite importante à cette méthode économétrique est le rang de $\widehat{H}$. Comme pour J dans les MCO, il est très probable que $\widehat{H}$ reste inversible si H ne l'est pas, mais la limite en probabilité de $\widehat{H}^{-1}$ n'est pas définie. Dans le cas où H serait nulle (par exemple si les données sont $i.i.d$), la convergence en $O\left(n^{\frac{-1}{2}}\right) = O\left(n_{obs}^{\frac{-1}{4}}\right)$ n'est pas garantie. Dans le cas $i.i.d$, cette convergence est en $O\left(n^{-1}\right) = O\left(n_{obs}^{\frac{-1}{2}}\right)$. Les intervalles asymptotiques sont donc fortement dépendants des hypothèses émises au début de l'étude économétrique, et rien (à ce jour) ne permet de détterminer si l'on est dans l'un ou l'autre des régimes de connvergence, ($n_{obs}^{-1/4}$ ou $n_{obs}^{-1/2}$), hormis l'intuition de l'économètre.

3 Implémentation en R

3.1 Généralités

Un répositoire pour l'utilisation de ces estimateurs en `R` se trouve à :
https://github.com/Yzeghal/Econometrics.git.
Il contient entre autres un `Readme.md` et les scripts `Sandbox_tester.R` et `Jointly Exchangeable Dissociated Array OLS.R`.
La configuration pour exécuter les fonctions dans leur version la plus simple est la suivante :

```
> sessionInfo()
R version 4.1.2 (2021-11-01)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 22621)

Matrix products: default

locale:
[1] LC_COLLATE=French_France.1252  LC_CTYPE=French_France.1252
[3] LC_MONETARY=French_France.1252 LC_NUMERIC=C
[5] LC_TIME=French_France.1252

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
[1] aod_1.3.2        matrixcalc_1.0-6

loaded via a namespace (and not attached):
[1] compiler_4.1.2 tools_4.1.2
```

Pour utiliser les estimateurs déjà présents dans des packages, il faut y ajouter `AER_1.2-10` et `R.utils_2.12.2`

Par souci de compréhension, les estimateurs des MCO avec estimateur de la variance asymptotique robuste à l'hétéroscédasticité ont été codés à nouveau, ainsi que l'estimateur des DMC. L'écart avec les résultats des fonctions `lm()` et `AER : :ivreg()` pour les $\widehat{\beta}$, `sandwich : :vcovHC(reg, type = "HC0")` pour la variance asymptotique des MCO et `lmtest : :wald.test()` pour le test de Wald est de l'ordre

de 10^{-11} pour des données simulées avec endogénéité et hétéroscédasticité. Un tel écart est attribuable au choix des opérations et aux approximations d'encodage.
Cependant, lorsque les données contiennent des `NA` distribués aléatoirement et en grand nombre, la convergence des estimateurs codés ici semble plus lente. De plus, certaines parties du code contiennent des boucles `for` qui sont beaucoup moins efficaces que des opérations vectorisées. Pour ces deux raisons, on préfèrera ne pas utiliser les estimateurs "maison". Lorsque le choix est possible, les fonctions ont un paramètre `built_in_reg` qui permet à l'utilisateur de choisir si les estimateurs à utiliser sont ceux codés dans le script ou dans les bibliothèques. Ces paramètres sont par défaut réglés sur `TRUE`.
Les deux fonctions de haut niveau destinées à l'utilisateur sont `OLS` et `IV_LS` , dont les appels et outputs sont illustrés ci-dessous :

```
OLS(X,Y,model="JEDA", hyp = 0, built_in_reg = TRUE)

>OLS_output_example
$coefs
     Beta_hat    Std_Err Student_t  p-values H0_hyp
[1,] 1.024731 0.71569157  1.431806 0.1522307      0
[2,] 1.999653 0.07511637 26.620737 0.0000000      0
[3,] 2.999165 0.07210449 41.594703 0.0000000      0

$var
          [,1]      [,2]       [,3]
[1,] 5122.1442 489.99880 -497.94448
[2,]  489.9988  56.42469  -53.47473
[3,] -497.9445 -53.47473   51.99057

$F
    chi2
14947462

$R2
[1] 0.9498505


IV_LS(G, Z, X, Y, hyp =0, built_in_reg = TRUE)
```

```
> IV_LS_output_example # 2 controls, 2 X endogenous , 2 exogenous
$SLS
$SLS$coefs
      Beta_hat    Std_Err Student_t     p-values H0_hyp
[1,] 0.9902366 0.17125433  5.782257 9.168481e-08      0
[2,] 1.9981640 0.11523234 17.340305 0.000000e+00      0
[3,] 3.0139547 0.04194214 71.859824 0.000000e+00      0
[4,] 0.9870646 0.04196550 23.520861 0.000000e+00      0

$SLS$var
             [,1]         [,2]         [,3]         [,4]
[1,]  2.932804528  1.851999089 -0.190323246  0.003563466
[2,]  1.851999089  1.327849299 -0.001697847 -0.125237183
[3,] -0.190323246 -0.001697847  0.175914305 -0.169749953
[4,]  0.003563466 -0.125237183 -0.169749953  0.176110290

$SLS$F
    chi2
15137066

$SLS$R2
[1] 0.5610715


$FLS
$FLS$R2
[1] 0.4824312 0.4970774

$FLS$F
[1] 4777.926 6182.389
```

Les arguments de ces fonctions sont :

- X : Variables explicatives (endogènes dans le cas des DMC, supposées exogènes dans le cas des OLS),
  Array de valeurs de taille (n,n,d_X), dont la première dimension doit contenir des 1 dans le cas des OLS
- Y : Variables expliquées,
  Array de valeurs de taille (n,n,1).

`G` : Variables de contrôle (DMC seulement),
Array de valeurs de taille (n,n,d_G), dont la première dimension doit contenir des `1` .

`Z` : Variables instrumentales (DMC seulement),
Array de valeurs de taille (n,n,d_Z).

`model` : Spécification du modèle (MCO seulement, les DMC sont toujours en mode 'JEDA').
-'JEDA' pour le cas des tableaux jointement échangeables dissociés,
-'BASIC' pour une variance asymptotique robuste à l'hétéroscadasticité,
-'NO ASVAR' pour n'avoir que l'estimateur des MCO

`hyp` : Vecteur des hypothèses pour les tests de student, de la même longueur que β. Fixé à 0 par défaut.

`built_in_reg` : Booléen, `FALSE` pour utiliser les estimateurs codés dans le script, `TRUE` pour utiliser les estimateurs codés dans les bibliothèques : `AER, Sandwich, lmtest` .

3.2 Génération de données de test

Le fichier `Sandbox_tester.R` contient des exemples de génération de données et d'utilisation de l'outil.
Pour tester ces fonctions, il faut générer des tableaux jointement échangeables de variables endogènes, d'erreurs hétéroscédastiques, d'instruments et de variables de contrôle. Pour ce faire on utilise la représentation d'Aldous Hoover Kallenberg en générant des tableaux $(U_i, U_j, U_{i,j})_{i,j}$, auxquels on applique une fonction. Le test des DMC a été réalisé sur des données générées par une fonction `g2` dont on peut modifier les coefficients.

```
eta_0 = matrix(c(1,2,3,1)) #constant coefficient included :(cst,G,X1,X2)
Beta_1 = matrix(c(10,10,5,7)) #(cst,Z1, Z2) #boost cst coef to max mean(x)
#and the OLS endogeneity error (on cst coef mainly)
Beta_2 = matrix(c(10,10,4,8)) #(cst,Z1, Z2)
n=100
Ui=runif(n,0,1)
Uij=matrix(runif(n^2,0,1),ncol=n)

g2<-function(UiUjUij){
  #UiUjUij is a vector c(Ui,Uj,Uij)
```

```
  #function that generates the Y, X, IVs Z and controls G
  U1=UiUjUij[1]
  U2=UiUjUij[2]
  U12=UiUjUij[3]
  p<-2*pi
  N1=cos(p*U12) #orthogonal to all other random variables
  N2=cos(3*p*U12)
  G = U1+3*cos(7*p*U12) #orthogonal to U2, U12 and their cos(2pi n . ) for n !=7 but
  eps1 = 30*(U1*U2/3+0.66)*N1
  eps2 = 30*(U1*U2/3+0.66)*N2
  eps3 = N1+N2 #Correlated to eps1, eps2 but not Z. Also heteroscedastic
  Z1=U1
  Z2=U2
  X1=t(Beta_1)%*%matrix(c(1,G,Z1,Z2))+eps1 #correlate G and X to have impact on G coe
  X2=t(Beta_2)%*%matrix(c(1,G,Z1,Z2))+eps2
  Y =t(Beta_0)%*%c(1,G,X1,X2)+eps3 #Y explained by X with endogeneity and heterosceda
  r=c(Y,1,G,Z1,Z2,X1,X2,eps1,eps2,eps3)
  return (r)
}


# data generation
#----
M=array(0,c(n,n,3))
M[,,1] = matrix(rep(Ui,n),c(n,n), byrow=FALSE)
M[,,2] = matrix(rep(Ui,n),c(n,n), byrow=TRUE) #test with the same i and not i,j
M[,,3] = Uij
A = apply(M,MARGIN = c(1,2),FUN = g2)
A=aperm(A,c(2,3,1))      #To have the dimensions we are used to work with

# for (i in 1:7){        #To dispatch NAs
#   A=dispatch_na(A,i)
# }

#A=diag_na(A)            #To put NAs in the diagonal
# Not having NAs in the diagonal does not change correlation estimators too much

#regression inputs
```

```
Y = A[,,1]
G = A[,,2:3]
Z = A[,,4:5]
X = A[,,6:7]
#----
#to estimate correlations
e1=as.vector(A[,,8])
e2=as.vector(A[,,9])
e3=as.vector(A[,,10])
x1=as.vector(X[,,1])
x2=as.vector(X[,,2])
z1=as.vector(Z[,,1])
z2=as.vector(Z[,,2])
k1=as.vector(G[,,2])
y =as.vector(Y)
```

On peut ensuite mesurer les corrélations empiriques :

```
> cor(e3,x1)
[1] 0.4090581
> cor(e3,x2)
[1] 0.4112386
> cor(e3,z1)
[1] -0.003598975
> cor(e3,z2)
[1] -0.001295877
> cor(e3,k1) #k1 is the control variable
[1] -0.009015544
> sd(e3)
[1] 0.9977797
```

L'indépendance des $(U_i)_i$ peut être admise par la construction `Ui=runif(n,0,1)` et l'orthogonalité est assurée par le fait que les fonctions $(t, \cos(2\pi t), \cos(4\pi t)...)$ forment une famille orthogonale sur $\mathcal{C}_{[0,1]}$.
L'orthogonalité entre Z_1 et ε_1 est obtenue par :

$$\mathbb{E}\left[(30(Z_1Z_2/3 + 0.66)N_1)Z_1\right] = \mathbb{E}\left[(30(Z_1Z_2/3 + 0.66)Z_1)\right] \underbrace{\mathbb{E}\left[N_1\right]}_{=0}$$

Car $N_1 \perp\!\!\!\perp Z_1, Z_2$. L'orthogonalité entre Z_1 et ε_2 s'obtient de la même manière, et l'orthogonalité entre Z_1 et ε_3 s'obtient par bilinéarité du produit scalaire (de même pour Z_2).

Remarquons qu'il n'est pas évident d'obtenir une corrélation empirique plus faible que 10^{-4} avec des données orthogonalement générées de cette manière, pour des raisons attribuées à la vitesse de convergence des estimateurs calculés par `corr()` . En effet, la convergence (15) implique que l'estimateur est bien consistant, mais pour $i, j \leq 10^3$, les écarts-types restent élevés car ils diminuent en $O(n^{-1/2}) = O(n_{obs}^{-1/4})$.

3.3 Tests sur données simulées

L'endogénéité est forte pour cet exemple, mais elle permet d'illustrer le choix d'utiliser `IV_LS` plutôt que `OLS` . Les résultats pour un jeu de données généré comme en section 3.2 sont les suivants :

```
GX=array(c(G,X),dim=c(dim(G)[1:2],dim(X)[3]+dim(G)[3])) #to compare with X
noIV<-OLS(GX,Y,model="JEDA")#runs OLS regression of Y on 1, G and X
noIV
regIV<-IV_LS(G,Z,X,Y,built_in_reg=TRUE) #runs the same 2SLS regression with IV
regIV
noIV>

$coefs
       Beta_hat      Std_Err  Student_t p-values H0_hyp
[1,] -0.3468731 0.0146636866  -23.65525        0      0
[2,]  1.1510514 0.0123297996   93.35524        0      0
[3,]  3.0421935 0.0006330104 4805.91409        0      0
[4,]  1.0419322 0.0006435448 1619.05148        0      0

$var
              [,1]          [,2]          [,3]          [,4]
[1,]  0.0215023704 -0.0082740164  3.564109e-04  3.776149e-04
[2,] -0.0082740164  0.0152023957 -7.505769e-04 -7.671067e-04
[3,]  0.0003564109 -0.0007505769  4.007021e-05  3.676179e-05
[4,]  0.0003776149 -0.0007671067  3.676179e-05  4.141499e-05

$F
     chi2
623552578

$R2
[1] 0.9999946
```

```
> regIV
$SLS
$SLS$coefs
     Beta_hat    Std_Err Student_t     p-values H0_hyp
[1,] 1.112157 0.14397021  7.724909 1.080558e-11      0
[2,] 2.050817 0.08494876 24.141811 0.000000e+00      0
[3,] 2.963523 0.05934690 49.935594 0.000000e+00      0
[4,] 1.030611 0.05360177 19.227178 0.000000e+00      0

$SLS$var
           [,1]       [,2]       [,3]       [,4]
[1,]  2.0727422  1.1974832 -0.6009633  0.4737104
[2,]  1.1974832  0.7216291 -0.3210197  0.2457585
[3,] -0.6009633 -0.3210197  0.3522054 -0.3157811
[4,]  0.4737104  0.2457585 -0.3157811  0.2873150

$SLS$F
    chi2
46836418

$SLS$R2
[1] 0.7603437


$FLS
$FLS$R2
[1] 0.6481838 0.6540882

$FLS$F
[1] 19373.51 18701.85
```

Ici, le R^2 de la régression de première étape a été artificiellement abaissé par le coefficient 30 devant les ε. On constate bien que l'estimateur des OLS ne converge pas vers β_0.

Références

[DDG22] Laurent Davezies, Xavier D'Haultfœuille et Yannick Guyonvarch. "The Marcinkiewicz–Zygmund law of large numbers for exchangeable arrays". In : *Statistics & Probability Letters* 188 (2022), p. 109536. issn : 0167-7152. doi : https://doi.org/10.1016/j.spl.2022.109536. url : https://www.sciencedirect.com/science/article/pii/S0167715222001031.

[FR98] Jeffrey A. Frankel et Andrew K. Rose. "The Endogeneity of the Optimum Currency Area Criteria". In : *The Economic Journal* 108.449 (1998), p. 1009-1025. issn : 00130133, 14680297. url : http://www.jstor.org/stable/2565665.

[Kal05] Olav Kallenberg. *"Probabilistic Symmetries and Invariance Principles"*. Springer New York, NY, 2005. isbn : 978-0-387-28861-1. doi : https://doi.org/10.1007/0-387-28861-9.