



HAL
open science

Modèles de régression avec variables muettes explicatives

Pietro Balestra

► **To cite this version:**

Pietro Balestra. Modèles de régression avec variables muettes explicatives. [Rapport de recherche] Institut de mathématiques économiques (IME). 1980, 33 p. hal-01526501

HAL Id: hal-01526501

<https://hal.science/hal-01526501>

Submitted on 23 May 2017

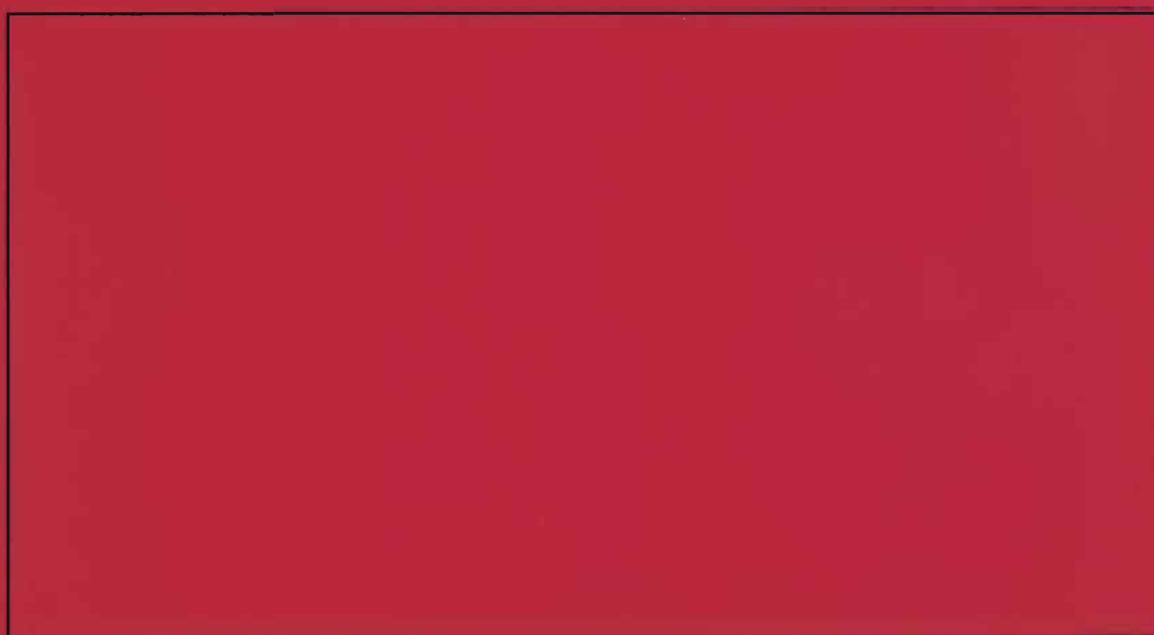
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

I.M.E.

EQUIPE DE RECHERCHE ASSOCIEE AU C.N.R.S.

DOCUMENT DE TRAVAIL



INSTITUT DE MATHEMATIQUES ECONOMIQUES

UNIVERSITE DE DIJON

FACULTE DE SCIENCE ECONOMIQUE ET DE GESTION

4, BOULEVARD GABRIEL — 21000 DIJON

N°39

MODELES DE REGRESSION AVEC
VARIABLES MUETTES EXPLICATIVES

Pietro BALESTRA

janvier 1980

Le but de cette Collection est de diffuser rapidement une première version de travaux poursuivis dans le cadre de l'I.M.E. afin de provoquer des discussions scientifiques. Les lecteurs désirant entrer en rapport avec un auteur sont priés d'écrire à l'adresse suivante:

Institut de Mathématiques Economiques - 4 Bd.Gabriel -
21000 Dijon (France).

TRAVAUX DEJA PUBLIES

- N°24 Pietro BALESTRA: Determinant and Inverse of a Sum of Matrices with Applications in Economics and Statistics (avril 1978)
- N°25 Bernard FUSTIER: Etude empirique sur la notion de région homogène (avril 1978)
- N°26 Claude PONSARD: On the Imprecision of Consumer's Spatial Preferences (avril 1978)
- N°27 Roland LANTNER: L'apport de la théorie des graphes aux représentations de l'espace économique (avril 1978)
- N°28 Emmanuel JOLLES: La théorie des sous-ensembles flous au service de la décision: deux exemples d'application (mai 1978)
- N°29 Michel PREVOT: Algorithme pour la résolution des systèmes flous (mai 1978)
- N°30 Bernard FUSTIER: Contribution à l'analyse spatiale de l'attraction imprécise (juin 1978)
- N°31 TRAN QUI Phuoc: Régionalisation de l'économie française par une méthode de taxinomie numérique floue (juin 1978)
- N°32 Louis De MESNARD: La dominance régionale et son imprécision, traitement dans le type général de structure (juin 1978)
- N°33 Max PINHAS: Investissement et taux d'intérêt. Un modèle stochastique d'analyse conjoncturelle (octobre 1978)
- N°34 Bernard FUSTIER, Bernard ROUGET: La nouvelle théorie du consommateur est-elle testable? (janvier 1979)
- N°35 Didier DUBOIS: Notes sur l'intérêt des sous-ensembles flous en analyse de l'attraction de points de vente (février 1979)
- N°36 Heinz SCHLEICHER, Equity Analysis of Public Investments: Pure and Mixed Game-Theoretic Solutions (April 1979)
- N°37 Jean JASKOLD GABSZEWICZ : Théories de la concurrence imparfaite : illustrations récentes de thèmes anciens (juin 1979).
- N° 38 Bernard FUSTIER : Contribution à l'étude d'un caractère statistique flou (janvier 1980).

Dans cette note nous considérons des modèles de régression dans lesquels on trouve, parmi les variables explicatives, des variables qualitatives. Le recours à des variables qualitatives explicatives est très fréquent dans les études économiques, comme en sociologie et dans d'autres disciplines, d'où l'importance d'une analyse systématique, sous l'angle méthodologique et opérationnel, du problème de l'estimation d'une régression en présence de variables qualitatives explicatives.

Dans le premier paragraphe, nous étudions le modèle de régression classique avec une seule variable muette à plusieurs attributs. On y verra que l'estimation du modèle peut se faire à l'aide d'une transformation très simple dont l'interprétation est fort intéressante. L'extension au cas de plusieurs variables muettes fait l'objet du deuxième paragraphe. On constatera dans ce paragraphe que la simplicité de la solution est perdue, lorsque l'on envisage des variables muettes multiples. Toutefois, un cas important de modèles à variables muettes multiples, le modèle à effets individuels et temporels, présente une solution facile et élégante. Ce modèle est discuté au paragraphe 3. Y-a-t-il d'autres cas de solution facile pour des modèles à variables muettes multiples? C'est la question que nous abordons dans le paragraphe 4, où une condition nécessaire et suffisante est donnée. Enfin, dans le dernier paragraphe, le problème des variables muettes est étudié dans le contexte de la régression généralisée, avec une application au modèle "seemingly unrelated regressions" de ZELLNER.

I - REGRESSION CLASSIQUE AVEC UNE SEULE VARIABLE MUETTE

Nous considérons une variable muette à s attributs, exhaustifs et mutuellement exclusifs, dans le cadre de la régression classique.

Quelques exemples suffisent pour montrer l'intérêt de ce type de variable qualitative.

Si la variable muette représente le sexe, les attributs sont deux, homme ou femme. Dans un modèle trimestriel, la variable muette représentant l'effet saisonnier aura quatre attributs, un pour chaque trimestre. Enfin, dans un modèle où la catégorie socio-professionnelle de l'individu est censée jouer un rôle, la variable muette aura autant d'attributs que de catégories socio-professionnelles.

Construisons s variables (vectorielles) auxiliaires L_1, L_2, \dots, L_s , une pour chaque attribut. Si la régression comporte n observations, chaque vecteur L_j , d'ordre n , est composé d'éléments nuls (pour les observations n'ayant pas l'attribut j) et d'éléments unitaires (pour les observations ayant l'attribut j).

Le modèle de régression classique avec variable muette s'écrira alors ainsi:

$$(1) \quad y = L_1 \alpha_1 + L_2 \alpha_2 + \dots + L_s \alpha_s + X \beta + \epsilon \\ = L \alpha + X \beta + \epsilon$$

où y est un vecteur $n \times 1$ d'observations sur une variable dépendante, X est une matrice d'ordre $n \times k$ d'observations sur k variables explicatives quantitatives (X ne doit pas contenir la colonne constante), L est la matrice $n \times s$ des variables auxiliaires associées à la variable muette, α est le vecteur $s \times 1$ des coefficients de la variable muette, β est le vecteur $k \times 1$ des coefficients des variables explicatives quantitatives et ϵ un vecteur $n \times 1$ d'erreurs. En régression classique, on admet que X est fixe et de rang complet $k < n+s$ et que $E(\epsilon) = 0$ et $E(\epsilon \epsilon') = \sigma^2 I$.

Pour chaque observation, une des variables auxiliaires prendra la valeur 1 et toutes les autres prendront la valeur zéro. Il s'en suit que les L_i sont mutuellement orthogonaux et que leur somme est égale au vecteur somme d'ordre n , noté S_n . On a donc:

- (2) $L_i' L_j = 0 \quad i \neq j$
 (3) $L_1 + L_2 + \dots + L_s = S_n \iff LS_s = S_n$
 (4) $L_j' L_j = n_j \iff L'L = \text{diag}(n_1, \dots, n_s)$

où n_j désigne le nombre d'observations ayant l'attribut j .

Evidemment

$$(5) n_j > 0 \quad \text{et} \quad \sum_{j=1}^s n_j = n .$$

On admet que la matrice $[L X]$ est de rang complet $s+k < n$.

Remarque 1

Dans la matrice X on ne peut pas avoir la colonne constante S_n , car elle serait linéairement dépendante des s colonnes L_j , ($LS_s = S_n$). On peut laisser la colonne S_n dans le modèle, seulement si l'on élimine un L_j .

Remarque 2

Si $s = 1$, $L = S_n$. Il s'agit alors du modèle de régression classique avec constante.

Le modèle de régression (1) peut être estimé directement par la méthode des moindres carrés ordinaires. Cette opération comporte l'inversion d'une matrice carrée d'ordre $k+s$. Lorsque le nombre s d'attributs est grand, il est intéressant de transformer le modèle afin d'inverser une matrice d'ordre k seulement. Quelle est cette transformation? Facilite-t-elle les calculs? C'est la question que nous nous posons ici.

Nous utilisons le Lemme suivant.

Lemme 1

Pour le modèle de régression

$$y = X_1 \beta_1 + X_2 \beta_2 + \epsilon$$

avec $[X_1 X_2]$ de rang complet, l'estimateur $\hat{\beta}_2$ de β_2 par les moindres carrés ordinaires peut s'exprimer ainsi:

$$\hat{\beta}_2 = [X_2' M_1 X_2]^{-1} X_2' M_1 y$$

$$\text{où } M_1 = I - X_1 (X_1' X_1)^{-1} X_1'$$

Pour $\hat{\beta}_1$ on obtient:

$$\hat{\beta}_1 = (X_1' X_1)^{-1} X_1' y - (X_1' X_1)^{-1} X_1' X_2 \hat{\beta}_2 .$$

En outre, les matrices des variances-covariances des deux estimateurs sont données, respectivement par:

$$\Omega_{\hat{\beta}_2} = \sigma^2 [X_2' M_1 X_2]^{-1}$$

$$\Omega_{\hat{\beta}_1} = \sigma^2 (X_1' X_1)^{-1} + (X_1' X_1)^{-1} X_1' X_1 (X_2' M_1 X_2)^{-1} X_2' X_1 (X_1' X_1)^{-1}$$

La preuve de ce résultat très connu peut se faire en exploitant les propriétés de l'inverse en forme partagée.

L'application directe du Lemme 1 au Modèle 1 donne

$$(6) \quad \hat{\beta} = (X' M_L X)^{-1} X' M_L y$$

où

$$(7) \quad M_L = I - L(L'L)^{-1} L' .$$

On se rappellera que la matrice M_L est idempotente d'ordre n et de rang $n-s$. Il en résulte que l'estimateur de β_2 peut s'obtenir en appliquant les moindres carrés ordinaires au modèle transformé suivant:

$$(8) \quad \tilde{y} = \tilde{X}\beta + \tilde{\epsilon}$$

où

$$\tilde{y} = M_L y \quad \text{et} \quad \tilde{X} = M_L X .$$

Que représente cette transformation? Considérons une variable transformée, disons \tilde{y}_i , $i=1, \dots, n$. Nous pouvons écrire

$$(9) \quad \tilde{y}_i = (e_i^n)' \tilde{y}$$

où e_i^n est le i -ème vecteur élémentaire d'ordre n . Ainsi:

$$\begin{aligned} \tilde{y}_i &= (e_i^n)' M_L y = (e_i^n)' (I - L(L'L)^{-1} L') y \\ &= y_i - (e_i^n)' L (L'L)^{-1} L' y . \end{aligned}$$

Supposons que l'observation i possède l'attribut j , $j=1, \dots, s$. Il vient:

$$(e_i^n)' L = (0 \ 0 \ \dots \ 1 \ \dots \ 0) = (e_j^s)'$$

$$(e_i^n)' L (L'L)^{-1} L' = (e_j^s)' (L'L)^{-1} L' = \frac{1}{n_j} (e_j^s)' L' = \frac{1}{n_j} L_j'$$

Au total donc:

$$(10) \quad \tilde{y}_i = y_i - \frac{1}{n_j} L_j' y .$$

La quantité $\frac{1}{n_j} L'_j y$ représente la moyenne des observations ayant le même attribut que l'observation i , c'est-à-dire la moyenne des observations possédant l'attribut j . Nous la désignons par le symbole \bar{y}_j . La même transformation s'applique aux variables explicatives. Ainsi:

$$(1) \quad \begin{aligned} \tilde{y}_i &= y_i - \bar{y}_j && \text{l'observation } i \text{ ayant l'attribut } j \\ \tilde{x}_{ir} &= x_{ir} - \bar{x}_{jr} && \text{" " " " " "} \end{aligned}$$

Par exemple, si la variable muette représente le sexe, pour un homme on enlève la valeur moyenne des hommes, et pour une femme on enlève la valeur moyenne des femmes.

Si l'on s'intéresse également aux coefficients α_j , par application directe du Lemme 1 on obtient:

$$(12) \quad \hat{\alpha} = (L'L)^{-1} L' y - (L'L)^{-1} L' X \hat{\beta}$$

ce qui donne pour chaque coefficient $\hat{\alpha}_j$

$$(13) \quad \hat{\alpha}_j = \bar{y}_j - \sum_{r=1}^k \bar{x}_{jr} \hat{\beta}_r .$$

La matrice des variances-covariances de $\hat{\beta}$ est:

$$(14) \quad \Omega_{\hat{\beta}} = \sigma^2 (X' M_L X)^{-1} = \sigma^2 (\tilde{X}' \tilde{X})^{-1}$$

alors que celle de $\hat{\alpha}$ est:

$$(15) \quad \Omega_{\hat{\alpha}} = \sigma^2 \left\{ (L'L)^{-1} + (L'L)^{-1} L' X (X' M_L X)^{-1} X' L (L'L)^{-1} \right\} .$$

Pour chaque coefficient $\hat{\alpha}_j$ on a la variance suivante

$$(16) \quad V(\hat{\alpha}_j) = \sigma^2 \left\{ \frac{1}{n_j} + [\bar{x}_{jr}]' (\tilde{X}' \tilde{X})^{-1} [\bar{x}_{jr}] \right\}$$

où

$$[\bar{x}_{jr}]' = [\bar{x}_{j1} \quad \bar{x}_{j2} \quad \dots \quad \bar{x}_{jk}] .$$

La formule (16) peut encore s'exprimer ainsi:

$$(17) \quad v(\hat{\alpha}_j) = \sigma^2 \frac{1}{n_j} + \sum_{r=1}^k \sum_{t=1}^k \bar{x}_{jr} \bar{x}_{jt} \text{cov}(\hat{\beta}_r, \hat{\beta}_t) .$$

Finalement, un estimateur centré de σ^2 est donné par

$$(18) \quad \hat{\sigma}^2 = \frac{\tilde{SS}}{n - (k+s)}$$

où \tilde{SS} est la somme des carrés des écarts du modèle transformé (8). Evidemment, \tilde{SS} est égal à SS , la somme des carrés des écarts du modèle (1).

Remarque 4

Lorsqu'on utilise l'ordinateur pour estimer le modèle transformé (8), le programme standard pour modèle de régression sans constante travaille avec $n-k$ degrés de liberté. Or, les degrés de liberté effectifs, en raison de la transformation, sont $n-k-s$. Il faudra donc ajuster toutes les variances et covariances par le facteur $\frac{n-k}{n-k-s}$.

Remarque 5

Il arrive que le programme de régression standard sur ordinateur ne prévoit pas la possibilité d'estimer un modèle sans constante.

Dans une telle situation on peut donner à l'ordinateur les données transformées correspondant au modèle (8). L'ordinateur travaillera avec la matrice $[S_n \tilde{X}]$ qui est de rang complet $k+1$ (S_n et \tilde{X} étant orthogonaux). Les résultats pour $\hat{\beta}$ sont rigoureusement les mêmes et l'estimation de la constante devrait donner zéro, à part les erreurs d'arrondi. Mais attention: les degrés de liberté utilisés par l'ordinateur sont dans ce cas $n-k-1$ et une correction appropriée est nécessaire pour les variances et covariances (facteur de correction $\frac{n-k-1}{n-k-s}$).

Remarque 6

Si l'on désire travailler explicitement avec les variables muettes en utilisant un programme d'ordinateur pour modèle avec constante, il est nécessaire d'éliminer un attribut, disons le dernier. Le modèle à estimer s'écrira:

$$y = L_1 \alpha_1 + \dots + L_{s-1} \alpha_{s-1} + S_n \alpha_0 + X \beta + \epsilon$$

et les α_j , $j=1, \dots, s-1$, représenteront alors l'écart par rapport au s -ème attribut. L'effet absolu pour les attributs est donné par $\alpha_j + \alpha_0$, pour les $s-1$ premiers attributs et par α_0 pour le s -ème. Avec cette interprétation, les résultats restent rigoureusement identiques.

Cela se voit de la manière suivante. Appelons L^{**} la matrice suivante:

$$L^{**} = [L_1 \ L_2 \ \dots \ L_{s-1} \ S_n]$$

L'estimation de β , par le Lemme 1, est

$$\hat{\beta} = (X' M_{L^{**}} X)^{-1} X' M_{L^{**}} y$$

Il suffit alors de montrer que $M_{L^{**}} = M_L$. Or, on voit immédiatement que

$$L^{**} = L A$$

où A est la matrice non-singulière suivante

$$A = \begin{bmatrix} I_{s-1} & S_{s-1} \\ 0 & 1 \end{bmatrix}$$

Il vient alors

$$\begin{aligned} M_{L^{**}} &= I - L^{**} (L^{**'} L^{**})^{-1} L^{**'} \\ &= I - L A (A' L' L A)^{-1} A' L \\ &= I - L A A^{-1} (L' L)^{-1} (A')^{-1} A' L \\ &= M_L . \end{aligned}$$

Désignons par $\hat{\epsilon}$ le vecteur d'erreurs calculées. En partant du modèle (1) de départ on a:

$$\begin{aligned} \hat{\epsilon} &= y - \hat{y} = y - L \hat{\alpha} - X \hat{\beta} \\ &= y - L (L' L)^{-1} L' y + L (L' L)^{-1} L' X \hat{\beta} - X \hat{\beta} = M_L y - M_L X \hat{\beta} \\ &= M_L (y - X \hat{\beta}) . \end{aligned}$$

Pour le modèle transformé (8), on a:

$$\hat{\epsilon} = \tilde{y} - \tilde{X} \hat{\beta} = M_L y - M_L X \hat{\beta} = M_L (y - X \hat{\beta}) .$$

Puisque $\hat{\beta}$ est le même dans les deux cas, les deux expressions pour $\hat{\epsilon}$ sont évidemment identiques. Il s'ensuit que dans le modèle à variable muette, comme dans le modèle classique de régression avec constante, la somme des erreurs calculées est automatiquement nulle. En effet:

$$S_n' \hat{\epsilon} = S_n' M_L (y - X \hat{\beta}) = 0$$

puisque

$$S_n' M_L = S_n' L M_L = 0 .$$

En plus, la somme des erreurs calculées pour toutes les observations de même attribut, disons j, est aussi automatiquement nulle car

$$L_j' \hat{\epsilon} = L_j' M_L (y - X \hat{\beta}) = 0 ,$$

Examinons maintenant un cas particulier, curieux et intéressant à la fois. Supposons qu'un attribut, disons le j -ème, est possédé par un seul individu (une seule observation), disons le i -ème. C'est le cas connu sous le nom d'événement unique.

La transformation pour cet individu donne

$$\tilde{y}_i = y_i - \bar{y}_j = y_i - y_i = 0$$

et il en va de même pour toutes les variables explicatives transformées (elles sont toutes nulles). La i -ème observation n'apporte donc rien à l'estimation du modèle. Elle disparaît par la transformation. On notera aussi, par le raisonnement fait plus haut, que l'erreur calculée pour cette observation est nulle ($L_j: \hat{\epsilon} = 0 \Rightarrow \hat{\epsilon}_j = 0$).

C'est comme si on éliminait la i -ème observation du modèle en même temps que le j -ème attribut. Les résultats resteraient rigoureusement les mêmes et le nombre de degrés de liberté serait inchangé.

Le coefficient α_j reste pourtant estimable. L'expression (13) dans la situation présente, donne

$$\hat{\alpha}_j = y_i - \sum_{r=1}^k x_{ir} \hat{\beta}_r .$$

Puisque $\hat{\beta}_r$ est le même que celui obtenu par les moindres carrés ordinaires en éliminant la i -ème observation (et j -ème attribut), l'estimateur $\hat{\alpha}_j$ est égal à l'erreur de prévision qui résulterait de la prévision de y_i sur la base des autres y_j et des valeurs de x_{ir} . Cette identité justifie l'utilisation d'une variable muette à événement unique pour tester si une observation supplémentaire appartient ou non au même univers. Plus précisément, ce test peut se faire soit en partant de l'erreur de prévision, soit en ajoutant l'observation supplémentaire avec une variable muette à événement unique. Dans ce dernier cas, on teste tout simplement la signification de α_j .

2 - EXTENSION A PLUSIEURS VARIABLES Muettes

Considérons le cas de deux variables muettes, l'une, comme auparavant, à s attributs et l'autre à t attributs. Considérons les $s+t$ variables auxiliaires L_1, L_2, \dots, L_s et T_1, T_2, \dots, T_t . Comme nous l'avons vu dans le paragraphe précédent, par la définition même de variable muette (exhaustive et mutuellement exclusive), on doit avoir:

$$(19) \quad \begin{cases} L_i' L_j = 0 & i \neq j \\ L_j' L_j = n_j & j = 1, \dots, s \\ L' L = \text{diag} (n_1, \dots, n_s) \\ L S_s = S_n \end{cases}$$

$$(20) \quad \begin{cases} T_i' T_j = 0 & i \neq j \\ T_j' T_j = m_j & j = 1, \dots, t \\ T' T = \text{diag} (m_1, \dots, m_t) \\ T S_t = S_n \end{cases}$$

où m_j , $0 < m_j \leq n$, $\sum_{j=1}^t m_j = n$, désigne le nombre d'observations ayant l'attribut associé à T_j .

A l'intérieur de chaque variable muette les variables auxiliaires sont mutuellement orthogonales, mais pas d'une variable muette à l'autre. On a:

$$(21) \quad \begin{cases} L_i' T_j = v_{ij} & i = 1, \dots, s \quad j = 1, \dots, t \\ L' T = V & V \text{ étant } s \times t \\ \sum_{j=1}^t v_{ij} = L_i' (T_1 + \dots + T_t) = L_i' S_n = n_i \\ \sum_{i=1}^s v_{ij} = (L_1' + \dots + L_s') T_j = S_n' T_j = m_j \end{cases}$$

On remarquera d'emblée que, même en éliminant la colonne constante, on ne peut pas inclure dans le modèle de régression les $s+t$ variables auxiliaires, car la matrice $[L \ T]$ ne peut pas être de rang complet. En effet, ses $s+t$ colonnes sont linéairement dépendantes, vu que:

$$[L \ T] \begin{bmatrix} S_s \\ -S_t \end{bmatrix} = S_n - S_n = 0.$$

Il s'avère donc nécessaire d'éliminer une variable auxiliaire. Cela peut être fait de manière arbitraire. Pour des raisons de symétrie, nous préférons exclure une variable auxiliaire pour chaque variable muette et inclure la colonne constante (Cette procédure s'applique également au cas de 3 variables muettes ou plus).

Le modèle de régression s'écrira alors:

$$(22) \quad y = L_1 \alpha_1 + \dots + L_{s-1} \alpha_{s-1} + T_1 \gamma_1 + \dots + T_{t-1} \gamma_{t-1} + S_n \alpha_0 + X\beta + \epsilon \\ = L^* \alpha^* + T^* \gamma^* + S_n \alpha_0 + X\beta + \epsilon$$

Les coefficients α_i , $i=1, \dots, s-1$ et γ_j , $j=1, \dots, t-1$, représentent alors l'écart par rapport à l'effet omis (représenté par α_0).

Le modèle (22) peut être estimé directement par les moindres carrés ordinaires. On pourrait également, comme au paragraphe précédent, utiliser la transformation qui élimine toutes les variables auxiliaires, y compris la constante. Désignons à cet effet la matrice Z suivante

$$(23) \quad Z = [L^* \ T^* \ S_n]$$

une matrice d'ordre $n \times (s+t-1)$ de rang complet. Le modèle transformé devient alors

$$(24) \quad \tilde{y} = \tilde{X} \beta + \tilde{\epsilon}, \quad \tilde{y} = M_Z y, \quad \tilde{X} = M_Z X, \quad M_Z = I - Z(Z'Z)^{-1} Z'$$

et l'estimateur de β est

$$(25) \quad \hat{\beta} = (X' M_Z X)^{-1} X' M_Z y.$$

Malheureusement la transformation M_Z est assez compliquée et son interprétation n'a pas la simplicité qu'on avait constaté dans le cas d'une seule variable muette. C'est la raison pour laquelle, sauf dans des situations particulières comme celles étudiées aux paragraphes 3 et 4, du point de vue opérationnel on se bornera à estimer directement le modèle (22).

Il est tout de même instructif d'analyser en détail un exemple, celui de deux variables muettes à deux attributs chacune ($s=t=2$). Supposons qu'on désire estimer l'effet d'un certain médicament. On dispose d'observations sur n individus, hommes et femmes, fumeurs et non-fumeurs. On pense que le sexe et le tabac jouent un rôle important. Pour vérifier cette hypothèse, on introduit dans le modèle deux variables muettes, une pour le sexe et l'autre pour la qualité d'être fumeur. Les quatre variables auxiliaires

sont: L_1 pour les hommes, L_2 pour les femmes, T_1 pour les fumeurs, T_2 pour les non-fumeurs. En accord avec les symboles énoncés plus haut, les caractéristiques des n individus peuvent être résumées dans le tableau suivant:

	Fumeurs	Non-fumeurs	
Hommes	v_{11}	v_{12}	n_1
Femmes	v_{21}	v_{22}	n_2
	m_1	m_2	n

La matrice Z prend alors la forme suivante

$$Z = [L_1 \ T_1 \ S_n]$$

et l'on calcule immédiatement $Z'Z$

$$Z'Z = \begin{bmatrix} n_1 & v_{11} & n_1 \\ v_{11} & m_1 & m_1 \\ n_1 & m_1 & n \end{bmatrix} = \begin{bmatrix} (v_{11}+v_{12}) & v_{11} & (v_{11}+v_{12}) \\ v_{11} & (v_{11}+v_{21}) & (v_{11}+v_{21}) \\ (v_{11}+v_{12}) & (v_{11}+v_{21}) & (v_{11}+v_{12}+v_{21}+v_{22}) \end{bmatrix}$$

Son déterminant et son inverse sont données respectivement par

$$\begin{aligned} d = |Z'Z| &= v_{11} v_{12} v_{21} + v_{11} v_{12} v_{22} + v_{11} v_{21} v_{22} + v_{12} v_{21} v_{22} \\ &= v_{11} v_{12} n_2 + v_{21} v_{22} n_1 \\ &= v_{11} v_{21} m_2 + v_{12} v_{22} m_1 \end{aligned}$$

$$(Z'Z)^{-1} = \frac{1}{d} \begin{bmatrix} (v_{11}+v_{21})(v_{12}+v_{22}) & v_{12}v_{21}-v_{22}v_{11} & -v_{12}(v_{11}+v_{21}) \\ v_{12}v_{21}-v_{22}v_{11} & (v_{11}+v_{12})(v_{21}+v_{22}) & -v_{21}(v_{11}+v_{12}) \\ -v_{12}(v_{11}+v_{21}) & -v_{21}(v_{11}+v_{12}) & v_{11}v_{12}+v_{11}v_{21}+v_{12}v_{21} \end{bmatrix}$$

Nous sommes en mesure maintenant de calculer la transformation $\tilde{y} = M_Z y$.
Considérons une variable transformée, disons la i -ème observation:

$$\begin{aligned} \tilde{y}_i &= (e_i^n)' M_Z y \\ &= y_i - (e_i^n)' Z (Z'Z)^{-1} Z' y \end{aligned}$$

Le vecteur $(e_i^n)' Z$ est un vecteur ligne à trois composantes donnant les valeurs prises par L_1 , T_1 et S_n pour le i -ème individu. Il peut prendre les quatre valeurs suivantes:

$$\begin{aligned}
R_1' &= [1 \ 1 \ 1] && \text{si homme et fumeur} && (L_1 \cap T_1) \\
R_2' &= [1 \ 0 \ 1] && \text{si homme et non-fumeur} && (L_1 \cap T_2) \\
R_3' &= [0 \ 1 \ 1] && \text{si femme et fumeur} && (L_2 \cap T_1) \\
R_4' &= [0 \ 0 \ 1] && \text{si femme et non-fumeur} && (L_2 \cap T_2)
\end{aligned}$$

L'individu i sera donc caractérisé par l'un des vecteurs ci-dessus, disons R_j , $j = 1, 2, 3, 4$. Désignons par

$$[r_{j1} \ r_{j2} \ r_{j3}]$$

le produit $R_j' (Z'Z)^{-1}$. Il vient alors:

$$\begin{aligned}
\tilde{y}_i &= y_i - R_j' (Z'Z)^{-1} Z' y \\
&= y_i - [r_{j1} \ r_{j2} \ r_{j3}] \begin{bmatrix} L_1' y \\ T_1' y \\ S_n' y \end{bmatrix} \\
&= y_i - r_{j1} L_1' y - r_{j2} T_1' y - r_{j3} S_n' y \\
&= y_i - r_{j1} n_1 \bar{y}_{L_1} - r_{j2} m_1 \bar{y}_{T_1} - r_{j3} n \bar{y}
\end{aligned}$$

où, comme d'habitude, \bar{y}_{L_1} désigne la moyenne des individus possédant l'attribut L_1 . On constate que la transformation a encore une interprétation assez simple: il suffit d'extraire certaines moyennes. Mais les coefficients r_{jk} sont fastidieux à calculer. Ils sont donnés dans le tableau suivant.

Coefficients r_{jk}

	k = 1	k = 2	k = 3
j = 1	$d^{-1} v_{21}(v_{21}+v_{22})$	$d^{-1} v_{12}(v_{21}+v_{22})$	$- d^{-1} v_{12}v_{21}$
j = 2	$d^{-1} v_{22}(v_{11}+v_{21})$	$- d^{-1} v_{11}(v_{21}+v_{22})$	$d^{-1} v_{11}v_{21}$
j = 3	$- d^{-1} v_{11}(v_{12}+v_{22})$	$d^{-1} v_{22}(v_{11}+v_{12})$	$d^{-1} v_{11}v_{12}$
j = 4	$- d^{-1} v_{12}(v_{11}+v_{21})$	$- d^{-1} v_{21}(v_{11}+v_{12})$	$d^{-1} (v_{11}v_{12}+v_{11}v_{21}+v_{12}v_{21})$

La transformation qu'on vient d'obtenir présente peut être un petit inconvénient. Elle suppose qu'on enlève toujours un multiple de \bar{y}_{L_1} et

\bar{y}_{T_1} . Or il se peut que l'individu i ne possède pas l'un ou l'autre (ou les deux) de ces attributs. Il paraît plus naturel d'extraire un multiple de la moyenne des attributs effectivement possédés par l'individu.

Il est facile de remédier à cet inconvénient, du fait que

$$L_1' y = (S_n' - L_2') y \text{ et que } T_1' y = (S_n' - T_1') y.$$

On a alors la transformation suivante

$$\begin{aligned} \tilde{y}_i &= y_i - d^{-1} a_{11} \bar{y}_{L_1} - d^{-1} a_{12} \bar{y}_{T_1} + d^{-1} a_{13} \bar{y} && \text{si homme et fumeur} \\ & && (L_1 \cap T_1) \\ &= y_i - d^{-1} a_{21} \bar{y}_{L_1} - d^{-1} a_{22} \bar{y}_{T_2} + d^{-1} a_{23} \bar{y} && \text{si homme et non-fumeur} \\ & && (L_1 \cap T_2) \\ &= y_i - d^{-1} a_{31} \bar{y}_{L_2} - d^{-1} a_{32} \bar{y}_{T_1} + d^{-1} a_{33} \bar{y} && \text{si femme et fumeur} \\ & && (L_2 \cap T_1) \\ &= y_i - d^{-1} a_{41} \bar{y}_{L_2} - d^{-1} a_{42} \bar{y}_{T_2} + d^{-1} a_{43} \bar{y} && \text{si femme et non-fumeur} \\ & && (L_2 \cap T_2) \end{aligned}$$

Les coefficients a_{ij} , dérivés des r_{ij} , sont donnés dans le tableau suivant.

Coefficients a_{ij}

$v_{21} m_2 n_1$	$v_{12} n_2 m_1$	$v_{12} v_{21} n$
$v_{22} m_1 n_1$	$v_{11} n_2 m_2$	$v_{11} v_{22} n$
$v_{11} m_2 n_2$	$v_{22} n_1 m_1$	$v_{11} v_{22} n$
$v_{12} m_1 n_2$	$v_{21} n_1 m_2$	$v_{12} v_{21} n$

Les calculs nécessaires à l'obtention des a_{ij} sont assez faciles, mais la transformation reste tout de même laborieuse. Dans certains cas particuliers, comme par exemple lorsque tous les v_{ij} sont égaux (c'est-à-dire $v_{ij} = 1/4 n$, $m_i = n_i = 1/2 n$), la transformation devient extrêmement simple, car $d^{-1} a_{ij} = 1$ pour tout i et pour tout j .

Dans ce cas, pour chaque individu, il suffit d'extraire les deux moyennes pour les deux attributs qu'il possède et d'ajouter la moyenne générale.

Le lecteur pourra vérifier qu'il en est également ainsi chaque fois que $v_{11}v_{22} = v_{12}v_{21}$. Nous reviendrons sur ce problème au paragraphe 4. Mais avant, dans le paragraphe qui suit, un autre cas particulier très intéressant est analysé.

3 - EFFETS INDIVIDUELS ET TEMPORELS

Un cas extrêmement intéressant se présente lorsque l'on dispose d'observations sur s individus (ou unités) à t moments différents dans le temps et que l'on suppose que dans le modèle de régression il y a un effet propre à l'individu et un effet propre en temps. Dans une telle situation, il y a au total st observations et le modèle de régression comporte deux variables muettes, la première à s caractères (pour l'effet individuel) et la seconde à t caractères (pour l'effet temporel).

Le modèle de régression complet, avec la notation du paragraphe précédent, s'écrira:

$$(26) \quad y = L^* \alpha^* + T^* \gamma^* + S_{st} \alpha_0 + X \beta + \epsilon$$

où L^* représente $(s-1)$ effets individuels et T^* représente $(t-1)$ effets temporels.

Il est de coutume dans ce type de modèle de munir les observations d'un double indice, le premier se référant à l'individu i , $i=1, \dots, s$, et le deuxième se référant à la période j , $j=1, \dots, t$. Les observations sont ordonnées en prenant d'abord les t observations pour le premier individu, puis les t observations pour le deuxième individu, etc. On désigne également par \bar{y}_i la moyenne du i -ème individu,

$$\bar{y}_i = \frac{1}{t} \sum_{j=1}^t y_{ij} = \frac{1}{t} L_i' y$$

par $y_{.j}$ la moyenne de la j -ème période,

$$\bar{y}_{.j} = \frac{1}{s} \sum_{i=1}^s y_{ij} = \frac{1}{s} T_j' y$$

et par $\bar{y}_{..}$ la moyenne globale,

$$\bar{y}_{..} = \frac{1}{st} \sum_{i=1}^s \sum_{j=1}^t y_{ij} = \frac{1}{st} S_{st}' y .$$

La variable L_i pour le i -ème effet individuel prend la forme

$$L_i = (e_i^s) \otimes S_t \Rightarrow L = I_s \otimes S_t, \quad L'L = t I_s, \quad L S_s = S_{st}$$

$$L' S_{st} = t S_s$$

et la variable T_j pour le j -ème effet temporel

$$T_j = S_s \otimes (e_j^t) \Rightarrow T = S_s \otimes I_t, T'T = sI_t, TS_t = S_{st}, T'S_{st} = sS_t.$$

On constate également que:

$$L'_i T_j = 1 \Rightarrow L'T = S_s S'_t.$$

A titre illustratif, pour 2 individus et quatre périodes, les variables auxiliaires prennent la forme explicite suivante

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} = [L_1 \ L_2 \ T_1 \ T_2 \ T_3 \ T_4]$$

Nous essayons maintenant de trouver la transformation M_Z qui élimine tous les effets individuels et temporels (y compris la constante). Conformément à (23), nous avons

$$Z = [L^* \ T^* \ S_{st}]$$

Puisque M_Z est invariant par rapport aux transformations non-singulières des colonnes de Z , nous pouvons prendre

$$Z_1 = [L \ T^*]$$

qui est obtenue de Z en enlevant de S_{st} la somme des colonnes de L^* (ce qui donne L_s) et en échangeant convenablement les colonnes. Mieux encore, on peut travailler avec

$$Z_2 = [L \ \tilde{T}]$$

où $\tilde{T} = T^* - \frac{1}{t} S_{st} S'_{t-1}$. Les colonnes de \tilde{T} sont obtenues en enlevant des colonnes correspondantes de T^* le vecteur $\frac{1}{t} S_{st}$ (c'est-à-dire $1/t$ fois la somme des colonnes de L). Cette représentation de la matrice Z offre l'avantage de l'orthogonalité entre L et T . En effet:

$$L'\tilde{T} = L'T^* - \frac{1}{t} L'S_{st} S'_{t-1} = S_s S'_{t-1} - \frac{1}{t} (t S_s) S'_{t-1} = 0.$$

Il vient alors:

$$Z_2 (Z_2' Z_2)^{-1} Z_2' = L (L'L)^{-1} L' + \tilde{T} (\tilde{T}'\tilde{T})^{-1} \tilde{T}' = \frac{1}{t} LL' + \tilde{T} (\tilde{T}'\tilde{T})^{-1} \tilde{T}'.$$

Il nous faut à présent calculer le dernier terme de l'expression ci-dessus. Commençons par le produit $\tilde{T}'\tilde{T}$:

$$\begin{aligned} \tilde{T}'\tilde{T} &= T^{\star'} T^{\star} - \frac{1}{t} T^{\star'} S_{st} S'_{t-1} - \frac{1}{t} S_{t-1} S'_{st} T^{\star} + \frac{1}{t^2} S_{t-1} S'_{st} S_{st} S'_{t-1} \\ &= s I_{t-1} - \frac{s}{t} S_{t-1} S'_{t-1} - \frac{s}{t} S_{t-1} S'_{t-1} + \frac{st}{t^2} S_{t-1} S'_{t-1} \\ &= s [I_{t-1} - \frac{1}{t} S_{t-1} S'_{t-1}] \end{aligned}$$

Son inverse est donné par:

$$(\tilde{T}'\tilde{T})^{-1} = \frac{1}{s} [I_{t-1} + S_{t-1} S'_{t-1}] .$$

Calculons maintenant $(\tilde{T}'\tilde{T})^{-1} \tilde{T}'$:

$$\begin{aligned} (\tilde{T}'\tilde{T})^{-1} \tilde{T}' &= \frac{1}{s} [I_{t-1} + S_{t-1} S'_{t-1}] [T^{\star'} - \frac{1}{t} S_{t-1} S'_{st}] \\ &= \frac{1}{s} [T^{\star'} + S_{t-1} S'_{t-1} T^{\star'} - \frac{1}{t} S_{t-1} S'_{st} - \frac{t-1}{t} S_{t-1} S'_{st}] \\ &= \frac{1}{s} [T^{\star'} + S_{t-1} (S'_{t-1} T^{\star'} - S'_{st})] . \end{aligned}$$

Puisque $T^{\star} S_{t-1} + T_s = S_{st}$, on a donc

$$(27) \quad (\tilde{T}'\tilde{T})^{-1} T' = \frac{1}{s} T^{\star'} - \frac{1}{s} S_{t-1} T'_s$$

Enfin

$$\begin{aligned} \tilde{T} (\tilde{T}'\tilde{T})^{-1} \tilde{T}' &= [T^{\star} - \frac{1}{t} S_{st} S'_{t-1}] [\frac{1}{s} T^{\star'} - \frac{1}{s} S_{t-1} T'_s] \\ &= \frac{1}{s} T^{\star} T^{\star'} - \frac{1}{st} S_{st} S'_{t-1} T^{\star'} - \frac{1}{s} T^{\star} S_{t-1} T'_s + \frac{t-1}{st} S_{st} T'_s \\ &= \frac{1}{s} T^{\star} T^{\star'} - \frac{1}{st} S_{st} (S'_{t-1} T^{\star'} + T'_s) + \frac{1}{s} (-T^{\star} S_{t-1} + S_{st}) T'_s \\ &= \frac{1}{s} T^{\star} T^{\star'} + \frac{1}{s} T_s T'_s - \frac{1}{st} S_{st} S'_{st} \\ &= \frac{1}{s} T T' - \frac{1}{st} S_{st} S'_{st} . \end{aligned}$$

Au total, la transformation M_Z est représentée par:

$$(28) \quad M_Z = I - \frac{1}{t} L L' - \frac{1}{s} T T' + \frac{1}{st} S_{st} S'_{st} .$$

Que représente cette transformation? Considérons la variable transformée \tilde{y}_{ij} pour l'individu i au temps j . Etant donné la manière dont nous avons ordonné les observations, cette observation se trouve dans la position $(i-1)t + j = r$ du vecteur y . Ainsi nous avons:

$$\begin{aligned}\tilde{y}_{ij} &= (e_r^{st})' \tilde{y} = (e_r^{st})' M_Z y \\ &= (e_r^{st})' (I - \frac{1}{t} L L' - \frac{1}{s} T T' + \frac{1}{st} S_{st} S_{st}') y \\ &= [(e_r^{st})' - \frac{1}{t} (e_i^s)' L' - \frac{1}{s} (e_j^t)' T' + \frac{1}{st} S_{st}'] y \\ &= y_{ij} - \frac{1}{t} L_i' y - \frac{1}{s} T_j' y + \frac{1}{st} S_{st}' y \\ &= y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..} .\end{aligned}$$

L'interprétation est alors facile. Pour transformer les variables (y compris les variables explicatives quantitatives) il suffit d'enlever la moyenne propre à l'individu, d'enlever la moyenne propre à la période et d'ajouter la moyenne générale.

Le coefficient β sera alors estimée tout simplement par

$$\hat{\beta} = (\tilde{X}' \tilde{X})^{-1} \tilde{X}' \tilde{y} .$$

Pour estimer les coefficients α^* , γ^* et α_0 (toujours en partant du modèle transformé) il convient de modifier l'écriture du modèle (26) afin de faire apparaître explicitement la matrice Z_2 . Ces modifications sont données ci-après:

$$\begin{aligned}(29) \quad y &= L^* \alpha^* + T^* \gamma^* + S_{st} \alpha_0 + X \beta + \epsilon \\ &= L^* \alpha^* + (\tilde{T} + \frac{1}{t} S_{st} S_{st}') \gamma^* + S_{st} \alpha_0 + X \beta + \epsilon \\ &= L^* \alpha^* + S_{st} (\alpha_0 + \frac{1}{t} S_{st}' \gamma^*) + \tilde{T} \gamma^* + X \beta + \epsilon \\ &= L^* \alpha^* + (L^* S_{s-1} + L_s) (\alpha_0 + \frac{1}{t} S_{st}' \gamma^*) + \tilde{T} \gamma^* + X \beta + \epsilon \\ &= L^* (\alpha^* + S_{s-1} (\alpha_0 + \frac{1}{t} S_{st}' \gamma^*)) + L_s (\alpha_0 + \frac{1}{t} S_{st}' \gamma^*) + \tilde{T} \gamma^* + X \beta + \epsilon \\ &= L^* g_1 + L_s g_2 + \tilde{T} g_3 + X \beta + \epsilon \\ &= Z_2 g + X \beta + \epsilon\end{aligned}$$

où g représente le vecteur des coefficients associés aux différentes colonnes de Z_2 . Par le Lemme 1 du paragraphe 1, nous obtenons

$$(30) \quad \hat{g} = (Z_2' Z_2)^{-1} Z_2' (y - X \hat{\beta})$$

et, en forme partagée, en utilisant (27)

$$(31) \quad \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \\ \hat{g}_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{t} L' \\ \frac{1}{t} L'_s \\ (T'T)^{-1} T' \end{bmatrix} (y - X \hat{\beta}) = \begin{bmatrix} \frac{1}{t} L' \\ \frac{1}{t} L'_s \\ \frac{1}{s} T' - \frac{1}{s} S_{t-1} T'_t \end{bmatrix} (y - X \hat{\beta})$$

Puisque $g_3 = \gamma^*$, la troisième ligne de l'expression (31) nous donne directement l'estimateur de γ^* . Pour chaque composante γ_j^* nous obtenons

$$(32) \quad \gamma_j^* = (\bar{y}_{.j} - \bar{y}_{.t}) - \sum_{r=1}^k (\bar{x}_{.jr} - \bar{x}_{.tr}) \hat{\beta}_r, \quad j = 1, \dots, t-1.$$

Pour α_0 , nous utilisons l'identité $g_2 = \alpha_0 + \frac{1}{t} S'_{t-1} \gamma^*$. Il vient:

$$(33) \quad \begin{aligned} \hat{\alpha}_0 &= \hat{g}_2 - \frac{1}{t} S'_{t-1} \hat{\gamma}^* \\ &= \hat{g}_2 - \left(\frac{1}{st} S'_{t-1} T'^* - \frac{t-1}{st} T'_t \right) (y - X \hat{\beta}) \\ &= \hat{g}_2 - \left(\frac{1}{st} (S'_{t-1} T'^* + T'_t) - \frac{1}{s} T'_t \right) (y - X \hat{\beta}) \\ &= \hat{g}_2 - \left(\frac{1}{st} S'_{st} - \frac{1}{st} S'_{st} \right) (y - X \hat{\beta}) = \left(\frac{1}{t} L'_s + \frac{1}{s} T'_t - \frac{1}{st} S'_{st} \right) (y - X \hat{\beta}) \\ &= (\bar{y}_{s.} + \bar{y}_{.t} - \bar{y}_{..}) - \sum_{r=1}^k (\bar{x}_{s-r} + \bar{x}_{.tr} - \bar{x}_{..r}) \hat{\beta}_r. \end{aligned}$$

Enfin, pour α^* nous utilisons l'identité $\alpha^* + S_{s-1} g_2 = g_1$, ce qui donne:

$$\hat{\alpha}^* = \left(\frac{1}{t} L'^* - \frac{1}{t} S_{s-1} L'_s \right) (y - X \hat{\beta}).$$

Pour chaque composante $\hat{\alpha}_i^*$ de ce vecteur, on a:

$$(34) \quad \hat{\alpha}_i^* = (\bar{y}_{i.} - \bar{y}_{s.}) - \sum_{r=1}^k (\bar{x}_{i.r} - \bar{x}_{s.r}) \hat{\beta}_r, \quad i=1, \dots, s-1.$$

La matrice des variances-covariances de $\hat{\beta}$ est évidemment donnée par

$$(35) \quad \Omega_{\hat{\beta}} = \sigma^2 (\tilde{X}' \tilde{X})^{-1}$$

qui est obtenue directement sur le modèle transformé. Pour les coefficients des variables muettes, nous obtenons d'abord, par le lemme 1, la matrice des variances-covariances de \hat{g} . Elle est la suivante

$$\begin{aligned}
(36) \quad \Omega_{\hat{g}} &= \sigma^2 (Z_2' Z_2)^{-1} + (Z_2' Z_2)^{-1} Z_2' X \Omega_{\hat{\beta}} X' Z_2 (Z_2' Z_2)^{-1} \\
&= \sigma^2 \begin{bmatrix} \frac{1}{t} I_{s-1} & 0 & 0 \\ 0 & \frac{1}{t} & 0 \\ 0 & 0 & \frac{1}{s} (I_{t-1} + S_{t-1} S_{t-1}') \end{bmatrix} \\
&\quad + \begin{bmatrix} \frac{1}{t} L' & & \\ \frac{1}{t} L'_s & & \\ \frac{1}{s} T'^* & -\frac{1}{s} S_{t-1} T'_t & \end{bmatrix} X \Omega_{\hat{\beta}} X' \begin{bmatrix} \frac{1}{t} L'^* & & \\ \frac{1}{t} L'_s & & \\ \frac{1}{t} T'^* & -\frac{1}{s} S_{t-1} T'_t & \end{bmatrix}'
\end{aligned}$$

Les coefficients $[\hat{\alpha}', \hat{\alpha}_0, \hat{\gamma}^*']$, dans cet ordre, sont une combinaison linéaire non-singulière de \hat{g} . En effet :

$$\begin{bmatrix} \hat{\alpha}^* \\ \hat{\alpha}_0 \\ \hat{\gamma}^* \end{bmatrix} = \begin{bmatrix} I_{s-1} & -S_{s-1} & 0 \\ 0 & 1 & -\frac{1}{t} S'_{t-1} \\ 0 & 0 & I_{t-1} \end{bmatrix} \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \\ \hat{g}_3 \end{bmatrix} = A \hat{g} .$$

Ainsi, nous obtenons

$$(37) \quad \Omega_{[\hat{\alpha}^*, \hat{\alpha}_0, \hat{\gamma}^*']} = A \Omega_{\hat{g}} A'$$

$$\begin{aligned}
&= \sigma^2 \begin{bmatrix} \frac{1}{t} (I_{s-1} + S_{s-1} S'_{s-1}) & -\frac{1}{t} S_{s-1} & 0 \\ -\frac{1}{t} S'_{s-1} & \frac{s+t-1}{st} & -\frac{1}{s} S'_{t-1} \\ 0 & -\frac{1}{s} S_{t-1} & \frac{1}{s} (I_{t-1} + S_{t-1} S'_{t-1}) \end{bmatrix} \\
&\quad + \begin{bmatrix} \frac{1}{t} L' & -\frac{1}{t} S_{s-1} L'_s \\ \frac{1}{t} L'_s + \frac{1}{t} T'_t - \frac{1}{st} S'_{st} \\ \frac{1}{s} T' & -\frac{1}{s} S_{t-1} T'_t \end{bmatrix} X \Omega_{\hat{\beta}} X' \begin{bmatrix} \frac{1}{t} L' & -\frac{1}{t} S_{s-1} L'_s \\ \frac{1}{t} L'_s + \frac{1}{s} T'_t - \frac{1}{st} S'_{st} \\ \frac{1}{s} T' & -\frac{1}{s} S_{t-1} T'_t \end{bmatrix}
\end{aligned}$$

Pour chaque coefficient individuel, les expressions suivantes peuvent en être déduites :

$$V(\hat{\alpha}_i^*) = \frac{2}{t} \sigma^2 + \sum_{r=1}^k \sum_{r'=1}^k (\bar{x}_{i.r} - \bar{x}_{s.r}) (\bar{x}_{i.r'} - \bar{x}_{s.r'}) \text{cov}(\hat{\beta}_r, \hat{\beta}_{r'})$$

$$V(\hat{\alpha}_0) = \frac{s+t-1}{st} \sigma^2 + \sum_{r=1}^k \sum_{r'=1}^k (\bar{x}_{s.r} + \bar{x}_{.tr} - \bar{x}_{..r}) (\bar{x}_{s.r'} + \bar{x}_{.tr'} - \bar{x}_{..r'}) \text{cov}(\hat{\beta}_r, \hat{\beta}_{r'})$$

$$V(\hat{\gamma}_j^*) = \frac{2}{s} \sigma^2 + \sum_{r=1}^k \sum_{r'=1}^k (\bar{x}_{.jr} - \bar{x}_{.tr}) (\bar{x}_{.jr'} - \bar{x}_{.tr'}) \text{cov}(\hat{\beta}_r, \hat{\beta}_{r'})$$

4 - CONDITION NECESSAIRE ET SUFFISANTE POUR UNE SOLUTION SIMPLE DANS
LE CAS DE VARIABLES Muettes MULTIPLES

Nous prenons ici comme modèle de référence le modèle (22) du paragraphe 2, dans lequel il y a deux variables muettes, la première à s attributs et la deuxième à t attributs (le nombre d'observations étant n).

Par solution simple nous entendons que le modèle transformé est obtenu à l'aide de la transformation qui consiste tout simplement à enlever pour chaque variable quantitative de la i -ème observation ($i=1, \dots, n$) les deux moyennes des deux attributs que cette observation possède et à ajouter la moyenne générale.

Le résultat fondamental est énoncé dans le théorème suivant.

Théorème 1

La condition nécessaire et suffisante pour que le modèle (22) possède une solution simple est

$$L'T = \frac{1}{n} N M'$$

où

$$N' = [n_1 \ n_2 \ \dots \ n_s] \quad \text{et} \quad M' = m_1 \ m_2 \ \dots \ m_t \ .$$

Avant de nous attaquer à la démonstration de ce théorème, montrons que cette condition est vérifiée dans les deux exemples traités aux paragraphes précédents.

Dans l'exemple du paragraphe 2 avec $s = t = 2$, on avait trouvé

$$(38) \quad L'T = \begin{vmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{vmatrix}$$

$$(39) \quad \frac{1}{n} N M' = \frac{1}{n} \begin{vmatrix} n_1 m_1 & n_1 m_2 \\ n_2 m_1 & n_2 m_2 \end{vmatrix}$$

$$= \frac{1}{v_{11}+v_{12}+v_{21}+v_{22}} \begin{vmatrix} (v_{11}+v_{12})(v_{11}+v_{21}) & (v_{11}+v_{12})(v_{12}+v_{22}) \\ (v_{21}+v_{22})(v_{11}+v_{21}) & (v_{21}+v_{22})(v_{12}+v_{22}) \end{vmatrix}$$

Pour que (38) soit égal à (39), il est nécessaire et suffisant que $v_{11}v_{22} = v_{12}v_{21}$, qui est la condition déjà donnée au paragraphe 2.

Pour le modèle du paragraphe 3, $L'T = S_s S_t'$, $n = st$, $N = t S_s$, $M = s S_t$ et la condition du théorème 1 est trivialement vérifiée.

La transformation du modèle (22) est donnée par la matrice M_Z . Nous pouvons prendre la matrice $Z = [L \ T^*]$. En utilisant la procédure d'inversion en forme partagée, on arrive facilement à l'expression suivante:

$$(40) \quad M_Z = M_L - M_L T^* (T^{*\prime} M_L T^*)^{-1} T^{*\prime} M_L$$

avec $M_L = I - L (L'L)^{-1} L'$.

La transformation simple, en revanche est donnée par la formule suivante:

$$(41) \quad Q = I - L(L'L)^{-1} L' - T(T'T)^{-1} T' + \frac{1}{n} S_n S_n'$$

$$= M_L - T(T'T)^{-1} T' + \frac{1}{n} S_n S_n'$$

Pour que les deux transformations soient identiques, il faut évidemment que

$$(42) \quad M_L T^* (T^{*\prime} M_L T^*)^{-1} T^{*\prime} M_L = T(T'T)^{-1} T' - \frac{1}{n} S_n S_n'$$

Nous sommes maintenant en mesure de prouver le théorème .

Nécessité de la condition

Supposons que (42) soit vrai. Pré-multiplions par L' (notant que $L'M_L = 0$). Il vient:

$$0 = L' T(T'T)^{-1} T' - \frac{1}{n} L' S_n S_n'$$

Post-multiplions par T :

$$0 = L'T - \frac{1}{n} L' S_n S_n' T$$

Or $L'S_n = [n_1 \ n_2 \ \dots \ n_s]' = N$ et $S_n' T = [m_1 \ m_2 \ \dots \ m_t] = M'$

On en déduit que

$$L'T = \frac{1}{n} N M'$$

Suffisance de la condition

Admettons maintenant que $L'T = \frac{1}{n} N M'$ et vérifions l'identité (42). Appelons M^* le vecteur M sans la dernière composante. De la condition $L'T = \frac{1}{n} N M'$ on tire évidemment que:

$$L' T^* = \frac{1}{n} N M^{*\prime}$$

Nous pouvons alors calculer $T^{\star'} M_L T^{\star}$:

$$\begin{aligned} T^{\star'} M_L T^{\star} &= T^{\star'} T^{\star} - \frac{1}{n} T^{\star'} L(L'L)^{-1} L' T^{\star} \\ &= T^{\star'} T^{\star} - \frac{1}{n} M^{\star} N' (L'L)^{-1} N M^{\star'} . \end{aligned}$$

Or $N' (L'L)^{-1} = [n_1 \ n_2 \ \dots \ n_s] \text{diag} \left(\frac{1}{n_1}, \dots, \frac{1}{n_s} \right) = S'_s$ et $S'_s N = \sum n_i = n$.

Il vient donc

$$T^{\star'} M_L T^{\star} = T^{\star'} T^{\star} - \frac{1}{n} M^{\star} M^{\star'} .$$

Par simple multiplication (notant que $(T^{\star'} T^{\star})^{-1} M^{\star} = S_{t-1}$) on vérifie que son inverse est donné par

$$(T^{\star'} M_L T^{\star})^{-1} = (T^{\star'} T^{\star})^{-1} + \frac{1}{m_t} S_{t-1} S'_{t-1} .$$

On trouve alors:

$$\begin{aligned} T^{\star} (T^{\star'} M_L T^{\star})^{-1} T^{\star'} &= T^{\star} (T^{\star'} T^{\star})^{-1} T^{\star'} + \frac{1}{m_t} T^{\star} S_{t-1} S'_{t-1} T^{\star'} \\ &= T^{\star} (T^{\star'} T^{\star})^{-1} T^{\star'} + \frac{1}{m_t} (S_n - T_t) (S_n - T_t)' \\ &= T^{\star} (T^{\star'} T^{\star})^{-1} T^{\star'} + \frac{1}{m_t} T_t T'_t - \frac{1}{m_t} T_s S'_n \\ &\quad - \frac{1}{m_t} S_n T'_t + \frac{1}{m_t} S_n S'_n \\ &= T (T'T)^{-1} T' - \frac{1}{m_t} T_t S'_n - \frac{1}{m_t} S_n T'_t + \frac{1}{m_t} S_n S'_n . \end{aligned}$$

Puisque $M_L S_n = 0$, on a:

$$M_L T^{\star} (T^{\star'} M_L T^{\star})^{-1} T^{\star'} M_L = M_L T (T'T)^{-1} T' M_L .$$

Pour développer la deuxième matrice ci-dessus, calculons d'abord $M_L T$:

$$\begin{aligned} M_L T &= T - L(L'L)^{-1} L' T = T - \frac{1}{n} L(L'L)^{-1} N M' \\ &= T - \frac{1}{n} L S_s M' = T - \frac{1}{n} S_n M' . \end{aligned}$$

Il vient alors:

$$\begin{aligned}
M_L T(T'T)^{-1} T' M_L &= (T - \frac{1}{n} S_n M') (T'T)^{-1} (T' - \frac{1}{n} M S'_n) \\
&= T(T'T)^{-1} T' - \frac{1}{n} S_n M' (T'T)^{-1} T' - \frac{1}{n} T(T'T)^{-1} M S'_n \\
&\quad + \frac{1}{n^2} S_n M' (T'T)^{-1} M S'_n \\
&= T(T'T)^{-1} T' - \frac{1}{n} S_n S'_t T' - \frac{1}{n} T S'_t S'_n + \frac{1}{n^2} S_n S'_t M S'_n \\
&= T(T'T)^{-1} T' - \frac{1}{n} S_n S'_n .
\end{aligned}$$

La condition (42) est donc vérifiée. Dans les passages ci-dessus, nous avons utilisé les propriétés suivantes:

$$M' (T'T)^{-1} = S'_t , \quad T S'_t + S_n \quad \text{et} \quad S'_t M = n .$$

La condition $L'T = \frac{1}{n} N M'$ peut évidemment s'écrire:

$$v_{ij} = \frac{n_i m_j}{n} , \quad \text{tout } i \text{ et } j ,$$

Or, v_{ij} par définition est un entier non négatif ($0 \leq v_{ij} \leq \min(n_i, m_j)$) alors que n_i et m_j sont des entiers strictement positifs ($0 < n_i < n$; $0 < m_j < n$). Cette constatation va nous permettre de tirer les deux corollaires suivants.

Corollaire 1

Si un v_{ij} (au moins) est nul, il n'y a pas de solution simple.

Corollaire 2

Si n est un nombre premier, il n'y a pas de solution simple.

5 - VARIABLES MUETTES DANS LA REGRESSION GENERALISEE

Nous avons vu au paragraphe 1 que, en présence d'une variable muette à s caractères exhaustifs et mutuellement exclusifs, avant d'appliquer les moindres carrés ordinaires on peut transformer le modèle et éliminer ainsi toutes les variables auxiliaires. La transformation, en outre, est extrêmement simple, car il suffit d'extraire de chaque observation la moyenne de toutes les observations ayant le même attribut.

Est-ce que cette procédure s'applique également en régression généralisée? En d'autres termes, est-ce que, avant d'appliquer les moindres carrés généralisés, on peut appliquer la transformation simple trouvée au paragraphe 1? C'est la question qui nous intéresse ici.

Formellement, nous étudions le modèle de régression généralisée suivant:

$$(43) \quad y = L \alpha + X \beta + \epsilon$$

où L représente s variables auxiliaires pour la variable muette, avec les propriétés énoncées en (2), (3) et (4), X est une matrice d'ordre $n \times k$ observations sur les variables explicatives (X est supposée de rang complet et ne doit pas contenir la colonne constante) et ϵ est un vecteur d'erreurs dont les propriétés sont

$$(44) \quad E(\epsilon) = 0 \quad E(\epsilon \epsilon') = \Omega_{\epsilon}$$

Ω_{ϵ} étant une matrice définie positive.

La transformation $M_L = I - L(L'L)^{-1} L'$ nous conduit au modèle transformé suivant

$$(45) \quad \tilde{y} = \tilde{X} \beta + \tilde{\epsilon}$$

avec $\tilde{y} = M_L y$ et $\tilde{X} = M_L X$. (Naturellement, comme dans la régression classique, le vecteur d'erreurs est également transformé. Mais on procède comme si le vecteur transformé possède les mêmes propriétés que le vecteur de départ).

La question avancée au début du paragraphe peut s'énoncer ainsi: sous quelle condition l'estimation par les moindres carrés généralisés du modèle (45) donne-t-elle le même estimateur pour β que l'estimation par les moindres carrés généralisés du modèle (43)? Appelons $\hat{\beta}$ l'estimateur obtenu en partant de (43) et β^* l'estimateur obtenu en partant de (45). La réponse à cette question est contenue dans le théorème suivant.

Théorème 2

La condition nécessaire et suffisante pour que l'estimateur des moindres carrés généralisés $\hat{\beta}$ soit égal à l'estimateur des moindres carrés généralisés β^* (avec même matrice des variances covariances Ω_ϵ) est

$$\Omega_\epsilon^{-1} L = L A$$

pour A non-singulière et arbitraire

Cette condition est évidemment équivalente à la condition suivante: il existe s vecteurs propres orthogonaux de Ω_ϵ^{-1} (ou, ce qui revient au même de Ω_ϵ) qui forment une base pour l'espace engendré par les s colonnes de L. Cette équivalence est montrée dans l'appendice. Nous préférons l'énoncé $\Omega_\epsilon^{-1} L = L A$, puisque, du point de vue opérationnel, il est plus simple.

Pour prouver le théorème, calculons d'abord $\hat{\beta}$. Puisque Ω_ϵ est définie positive, il existe une matrice non-singulière P telle que

$$(46) \quad P \Omega_\epsilon P' = I .$$

c'est-à-dire

$$(47) \quad P' P = \Omega_\epsilon^{-1} .$$

Appliquons la transformation P au modèle (43). Il vient

$$(48) \quad P y = P L \alpha + P X \beta + P \epsilon$$

Le vecteur $P \epsilon$ possède maintenant une matrice des variances covariances unitaire et les moindres carrés ordinaires sont appropriés. Grâce au Lemme 1, nous obtenons directement

$$(49) \quad \hat{\beta} = (X' P' M_{PL} P X)^{-1} X' P' M_{PL} P y$$

où

$$(50) \quad M_{PL} = I - PL (L' P' P L)^{-1} L' P' = I - PL (L' \Omega_\epsilon^{-1} L)^{-1} L' P' .$$

D'autre part, l'estimation par les moindres carrés généralisés (avec matrice des variances covariances Ω_ϵ) du modèle transformé (45) donne

$$(51) \quad \begin{aligned} \beta^* &= (X' \Omega_\epsilon^{-1} X)^{-1} X' \Omega_\epsilon^{-1} y \\ &= (X' M_L \Omega_\epsilon^{-1} M_L X)^{-1} X' M_L \Omega_\epsilon^{-1} M_L y . \end{aligned}$$

Pour que $\hat{\beta}$ soit égal à β^* , pour toute valeur de la matrice X, il est nécessaire et suffisant que

$$(52) \quad P' M_{PL} P = M_L \Omega_\epsilon^{-1} M_L .$$

Notons que, par la propriété d'idempotence de M_{PL} , $P' M_{PL} P L = 0$ de sorte que $(P' M_{PL} P) M_L = P' M_{PL} P$. Egalement, $L' (P' M_{PL} P) = 0$. Nous pouvons donc écrire

$$(53) \quad \begin{aligned} P' M_{PL} P &= M_L (P' M_{PL} P) M_L \\ &= M_L (P'P - P'PL (L' \Omega_\epsilon^{-1} L)^{-1} L' P'P) M_L \\ &= M_L (\Omega_\epsilon^{-1} - \Omega_\epsilon^{-1} L (L' \Omega_\epsilon^{-1} L)^{-1} L' \Omega_\epsilon^{-1}) M_L \\ &= M_L \Omega_\epsilon^{-1} M_L - M_L \Omega_\epsilon^{-1} L (L' \Omega_\epsilon^{-1} L)^{-1} L' \Omega_\epsilon^{-1} M_L . \end{aligned}$$

La condition (52) peut maintenant s'exprimer ainsi

$$(54) \quad M_L \Omega_\epsilon^{-1} L (L' \Omega_\epsilon^{-1} L)^{-1} L' \Omega_\epsilon^{-1} M_L = 0 .$$

Puisque la matrice $(L' \Omega_\epsilon^{-1} L)^{-1}$ est définie positive, l'égalité ci-dessus est satisfaite si et seulement si

$$(55) \quad M_L \Omega_\epsilon^{-1} L = 0 .$$

La matrice M_L est idempotente et de rang s . Il s'ensuit que le système homogène $M_L h = 0$ admet exactement s solutions linéairement indépendantes. Puisque $M_L L = 0$, L étant de rang s , les s colonnes de L représentent s solutions indépendantes du système homogène. Toute autre solution peut être exprimée par une combinaison linéaire des colonnes de L . Il s'ensuit que (55) est vérifiée si et seulement si les s colonnes de $\Omega_\epsilon^{-1} L$ sont s combinaisons linéaires indépendantes de L , c'est-à-dire si et seulement si

$$(56) \quad \Omega_\epsilon^{-1} L = L A$$

pour A non-singulière. Le théorème est ainsi prouvé.

Si la condition (56) est vérifiée, la matrice des variances covariances de β^* se trouve directement en partant du modèle transformé (45)

On a :

$$(57) \quad \Omega_{\beta^*} = (\tilde{X}' \Omega_\epsilon^{-1} \tilde{X})^{-1} .$$

En ce qui concerne les coefficients de la variable muette, par application du Lemme 1, on a :

$$\begin{aligned}\hat{\alpha} &= (L' P' P L)^{-1} L' P' P (y - X \beta^*) \\ &= (L' \Omega_\epsilon^{-1} L)^{-1} L' \Omega_\epsilon^{-1} (y - X \beta^*) .\end{aligned}$$

Or, si la condition (56) est vérifiée, cette formule se simplifie considérablement. Calculons d'abord l'expression suivante :

$$(L' \Omega_\epsilon^{-1} L)^{-1} L' \Omega_\epsilon^{-1} = (A' L' L)^{-1} A' L' = (L'L)^{-1} L'$$

Il vient alors

$$(58) \quad \hat{\alpha} = (L'L)^{-1} L' (y - X \beta^*)$$

ce qui donne, pour chaque coefficient individuel

$$(59) \quad \hat{\alpha}_j = y_j - \sum_{r=1}^k x_{jr} \beta_r^* .$$

La matrice des variances covariances de $\hat{\alpha}$, toujours en invoquant le Lemme 1, est donnée par

$$\begin{aligned}(60) \quad \Omega_{\hat{\alpha}} &= (L' \Omega_\epsilon^{-1} L)^{-1} + (L' \Omega_\epsilon^{-1} L)^{-1} L' \Omega_\epsilon^{-1} X(X' P' M_{PL} P X)^{-1} \\ &\quad X' \Omega_\epsilon^{-1} L(L' \Omega_\epsilon^{-1} L)^{-1} \\ &= (L' \Omega_\epsilon^{-1} L)^{-1} + (L'L)^{-1} L' X (X' \Omega_\epsilon^{-1} X)^{-1} X' L (L'L)^{-1} \\ &= A^{-1} (L'L)^{-1} + (L'L)^{-1} L' X \Omega_{\beta^*} X L (L'L)^{-1} ,\end{aligned}$$

Pour chaque coefficient individuel, cette expression donne :

$$(61) \quad V(\hat{\alpha}_j) = \frac{1}{n_j} a^{jj} + \sum_{r=1}^k \sum_{r'=1}^k x_{jr} x_{jr'} \text{cov}(\beta_r^*, \beta_{r'}^*)$$

où a^{jj} est l'élément dans la position (j,j) de A^{-1} .

Une application intéressante de ce résultat au modèle "Seemingly unrelated regressions" de ZELLNER (1) est présentée ci-après. Dans ce modèle on a s régressions de t observations chacune. Si les régressions contiennent la constante, pour la i-ème régression on a :

(1) ZELLNER A. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, Journal of the American Statistical Association, vol.57 (1962), pp.348-68.

$$(62) \quad y_i = S_t \alpha_i + X_i \beta_i + \epsilon_i$$

S_t étant le vecteur somme d'ordre t , y_i étant le vecteur d'ordre t d'observations concernant la variable dépendante, X_i étant une matrice txk_i d'observations concernant les variables explicatives et ϵ_i étant un vecteur d'erreurs dont les propriétés sont énoncées plus loin.

Le modèle complet peut s'écrire

$$(63) \quad \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_s \end{bmatrix} = \begin{bmatrix} S_t & & & \\ & S_t & & \\ & & \ddots & \\ & & & S_t \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_s \end{bmatrix} + \begin{bmatrix} X_1 & & & \\ & X_2 & & \\ & & \ddots & \\ & & & X_s \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_s \end{bmatrix}$$

et en forme compacte

$$(64) \quad y = L \alpha + X \beta + \epsilon .$$

On notera que $L = I_s \otimes S_t$ représente une variable muette à s caractères. Le vecteur α est le vecteur des coefficients de la variable muette. Dans ce type de modèle, le vecteur d'erreurs est caractérisé par une espérance mathématique nulle, $E(\epsilon) = 0$, et par la matrice des variances covariances suivante

$$(65) \quad \Omega_\epsilon = \Omega \otimes I_t$$

où $\Omega = [\sigma_{ij}]$ est une matrice définie positive d'ordre s .

Calculons $\Omega_\epsilon^{-1} L$:

$$\begin{aligned} \Omega_\epsilon^{-1} L &= (\Omega^{-1} \otimes I_t)(I_s \otimes S_t) = (\Omega^{-1} \otimes S_t) \\ &= (I_s \otimes S_t)(\Omega^{-1} \otimes 1) = (I_s \otimes S_t) \Omega^{-1} \\ &= L \Omega^{-1} \end{aligned}$$

On voit que la condition du théorème 2 est satisfaite avec $A = \Omega^{-1}$. Cela signifie que dans le modèle de ZELLNER, l'estimation par les moindres carrés généralisés peut se faire en éliminant pour chaque régression la colonne constante et en exprimant toutes les autres variables en déviation par rapport à la moyenne de la régression concernée.

Il est intéressant de noter que pour établir le théorème 2 nous n'avons pas utilisé les propriétés spécifiques de la variable muette L. Le résultat est donc valable en général, quelle que soit L, pourvu que la matrice $[L X]$ soit de rang complet $s \times k$. La seule différence est que la matrice de la transformation M_L n'aura pas une structure particulièrement simple.

Cette constatation va nous permettre d'étendre le résultat du théorème 2 au cas de deux variables muettes, la première à s attributs et la deuxième à t attributs.

Reprenons le modèle (22) avec, cette fois, $E(\epsilon\epsilon') = \Omega_\epsilon$ et demandons nous si, avant d'appliquer les moindres carrés généralisés, on peut transformer le modèle par la transformation simple représentée par la matrice Q, voir (41). La réponse est contenue dans le théorème suivant.

Théorème 3

La condition nécessaire et suffisante pour que le modèle (22) avec $E(\epsilon\epsilon') = \Omega_\epsilon$ possède une solution simple par les moindres carrés généralisés est

$$Q \Omega_\epsilon^{-1} Z = 0 .$$

Cette condition est équivalente aux deux conditions suivantes

$$(i) \quad L'T = \frac{1}{n} N M'$$

$$(ii) \quad \Omega_\epsilon^{-1} Z = Z B$$

pour B non-singulière.

La démonstration de ce théorème est la même que celle du théorème 2. Il suffit de remplacer L par Z et M_L par Q, en notant que $P' M_{PZ} P = Q(P' M_{PZ} P) Q$. On obtient alors directement l'analogue de la condition (55), c'est-à-dire $Q \Omega_\epsilon^{-1} Z = 0$.

Pour montrer l'équivalence de la condition $Q \Omega_\epsilon^{-1} Z = 0$ et des deux conditions (i) et (ii) du théorème 3 procédons par étapes.

Suffisance

Si (ii) est vrai, alors

$$Q \Omega_\epsilon^{-1} Z = Q Z B .$$

Maintenant, si (i) est également vrai, nous avons établi au paragraphe 4 que $Q = M_Z$. Il vient alors

$$Q \Omega_{\epsilon}^{-1} Z = Q Z B = M_Z Z B = 0 .$$

La condition $Q \Omega_{\epsilon}^{-1} Z = 0$ est donc vérifiée .

Nécessité

Puisque $Z = [L T^*]$, la condition $Q \Omega_{\epsilon}^{-1} Z = 0$ implique $Q \Omega_{\epsilon}^{-1} L = 0$. Pré-multiplions cette dernière expression par T' et développons:

$$T' Q \Omega_{\epsilon}^{-1} L = 0$$

$$T' \Omega_{\epsilon}^{-1} L - T' L(L'L)^{-1} L' \Omega_{\epsilon}^{-1} L - T'T (T'T)^{-1} T' \Omega_{\epsilon}^{-1} L \\ + \frac{1}{n} T' S_n S_n' \Omega_{\epsilon}^{-1} L = 0$$

$$- T'L (L'L)^{-1} L' \Omega_{\epsilon}^{-1} L + \frac{1}{n} M S_S' L' \Omega_{\epsilon}^{-1} L = 0 \quad (T'S_n = M, S_n' = S_S' L')$$

$$T'L (L'L)^{-1} = \frac{1}{n} M S_S' \quad (L' \Omega_{\epsilon}^{-1} L \text{ non-singulière})$$

$$T'L = \frac{1}{n} M S_S' L'L$$

$$T'L = \frac{1}{n} M S_n' L$$

$$T'L = \frac{1}{n} M N' .$$

La condition (i) est donc nécessaire. Elle implique également, nous l'avons vu, $Q = M_Z$. Donc, de la condition $Q \Omega_{\epsilon}^{-1} Z = 0$ nous tirons:

$$M_Z \Omega_{\epsilon}^{-1} Z = 0 .$$

Or cette dernière condition est satisfaite si et seulement si $\Omega_{\epsilon}^{-1} Z = Z B$, c'est-à-dire si (ii) est vrai.

APPENDICE

Nous montrons dans cette appendice que la condition $\Omega_\epsilon^{-1} L = L A$ est équivalente à la condition $L = C_1 B$. Les matrices Ω_ϵ^{-1} , L et A sont définies au paragraphe 5. La matrice B est une matrice non-singulière d'ordre k et C_1 représente une matrice $n \times k$ de vecteurs propres orthogonaux de Ω_ϵ^{-1} , c'est-à-dire $\Omega_\epsilon^{-1} C_1 = C_1 \Lambda_1$ où Λ_1 est une matrice diagonale contenant k valeurs propres de Ω_ϵ^{-1} .

Première étape : si $L = C_1 B$, alors $\Omega_\epsilon^{-1} L = L A$.

Preuve: $\Omega_\epsilon^{-1} L = \Omega_\epsilon^{-1} C_1 B = C_1 \Lambda_1 B = C_1 B B^{-1} \Lambda_1 B = L A$,
 $A = B^{-1} \Lambda_1 B$.

Deuxième étape : si $\Omega_\epsilon^{-1} L = L A$, alors $L = C_1 B$.

Preuve: $\Omega_\epsilon^{-1} L = L A$
 $L' \Omega_\epsilon^{-1} L = L' L A$
 $A = (L' L)^{-1} (L' \Omega_\epsilon^{-1} L)$.

Donc: $\Omega_\epsilon^{-1} L = L (L' L)^{-1} (L' \Omega_\epsilon^{-1} L)$.

Maintenant: $(L' L)^{-1}$ est définie positive. Il existe donc une matrice non singulière P telle que $P' P = (L' L)^{-1}$.

On a alors:

$$\Omega_\epsilon^{-1} L = L P' P L' \Omega_\epsilon^{-1} L$$

$$\Omega_\epsilon^{-1} L P' = L P' P L' \Omega_\epsilon^{-1} L P'.$$

Or, la matrice $P L' \Omega_\epsilon^{-1} L P'$ est définie positive. Il existe donc une matrice orthogonale Q telle que

$$P L' \Omega_\epsilon^{-1} L P' = Q \Lambda Q'$$

où Λ est la matrice diagonale (non singulière) des valeurs propres. Il vient donc

$$\Omega_\epsilon^{-1} L P' = L P' Q \Lambda Q'$$

$$\Omega_\epsilon^{-1} L P' Q = L P' Q \Lambda$$

$$\Omega_\epsilon^{-1} C_1 = C_1 \Lambda, \quad C_1 = L P' Q.$$

Les colonnes de C_1 sont orthogonales, $C_1' C_1 = Q' P L' L P' Q = Q' P(P'P)^{-1} P' Q$
 $= Q' Q = I$,

et représentent k vecteurs propres orthogonaux de Ω_e^{-1} . Finalement:

$$L P' Q = C_1$$

$$L = C_1 Q' (P')^{-1}$$

$$L = C_1 B \quad , \quad B = Q' (P')^{-1}.$$