Les régressions linéaires multiples

Définition

On souhaite connaitre la relation entre une variables d'intérêt y et plusieurs variables x_1, x_2, ..., x_K dans lapopulation.

La relation la plus simple est linéaire:

$$y = a + b_1 x_1 + b_2 x_2 + ... + b_K x_K$$

Voici la régression multiple de y sur $x_1, ..., x_K$.

Exemple: l'équation des salaires de Mincer

- Data = U.S. CPS (Current Population Survey)
- la variable dépendante *y* est le log du revenu.
- Les régresseurs sont les déterminants du capital humain. x1 est le nombre d'années d'études, x2 l'expérience, x3 l'expérience au carré etc …

Pourquoi ajouter des variables

- La théorie l'impose
  - ex: Cobb-Douglas : le K et le L dans la fonction de production
- La forme fonctionnelle l'impose
  - ex: Courbe de Kuznets : la forme en « U » inversé
- Notre intuition nous le suggère
  - ex: La consommation médicale augmente avec l'âge (dégradation de l'état de santé) mais aussi avec le salaire (effet du reste à charge)
- Analyse ceteris paribus, : on contrôle des effets des autres variables … modèles « ajustés »

Introduire une nouvelle variable explicative

- Rappel régression par MCO :

$$y = \beta 1 x + \beta 0 + u$$

  - S'interprète comme une décomposition de la variance : SCT=SCE+SCR
  - C'est une ANOVA = ANalysis Of Variance
  - On rajoute simplement des x et des β pour tenir compte des autres facteurs
  - On cherche à isoler les effets de x1 et de x2 de leur influence réciproque sur y
  - On cherche l'effet de x1 sur la variance de y purgée de l'effet de x2

La théorie est exactement la même que dans le modèle simple:

- soit un échantillon d'observations $\{(y_i, x_{1i}, ..., x_{Ki}), i = 1, ..., N\}$.
- On note l'estimateur OLS $\hat{a}, \hat{b}_1, ..., \hat{b}_K$ et les prédictions :

$$\hat{y}_i = a + \hat{b}_1 x_{1i} + \hat{b}_2 x_{2i} + ... + \hat{b}_K x_{Ki}$$
$$\hat{u}_i = y_i - \hat{y}_i$$

L'OLS assignes des valeurs à chaque coefficient de telle sorte à minimiser la somme des carrés résiduelle : $SSR = \sum_{i=1}^{N} \hat{u}_i^2$

Les équations normales impliquent :

- Les résidus on une moyenne de zéro : E(Ûi) =0
- La covariance entre les résidu et chaque régresseur est nulle:

$$Cov_N(x_{ki}, \hat{u}_i) = \frac{1}{N} \sum_{i=1}^{N} x_{ki} \hat{u}_i = 0, \text{ for } k = 1, ..., K.$$

OLS derivation 1

Formellement le problème est noté :

$$\min_{(a,b)} \left\{ SSR(a, b_1, ..., b_K) \equiv \sum_{i=1}^{N} (y_i - a - b_1 x_{1i} - b_2 x_{2i} - ... - b_K x_{Ki})^2 \right\}$$

Les conditions de premier ordre (optimalité) :

$$\frac{\partial SSR(\widehat{a}, \widehat{b}_1, ..., \widehat{b}_K)}{\partial a} = -2 \sum_{i=1}^{N} (y_i - \widehat{a} - \widehat{b}_1 x_{1i} - \widehat{b}_2 x_{2i} - ... - \widehat{b}_K x_{Ki}) = 0$$

$$\frac{\partial SSR(\widehat{a}, \widehat{b}_1, ..., \widehat{b}_K)}{\partial b_1} = -2 \sum_{i=1}^{N} x_{1i} (y_i - \widehat{a} - \widehat{b}_1 x_{1i} - \widehat{b}_2 x_{2i} - ... - \widehat{b}_K x_{Ki}) = 0$$

$$...$$

$$\frac{\partial SSR(\widehat{a}, \widehat{b}_1, ..., \widehat{b}_K)}{\partial b_K} = -2 \sum_{i=1}^{N} x_{Ki} (y_i - \widehat{a} - \widehat{b}_1 x_{1i} - \widehat{b}_2 x_{2i} - ... - \widehat{b}_K x_{Ki}) = 0$$

OLS derivation 2

donc,

$$\frac{1}{N}\sum_{i=1}^{N}\widehat{u}_i = 0, \quad \frac{1}{N}\sum_{i=1}^{N}x_{ki}\widehat{u}_i = 0, k = 1, ..., K$$

On note que :

$$\text{Cov}_N(x_{ki}, \widehat{u}_i) = \frac{1}{N}\sum_{i=1}^{N}x_{ki}\widehat{u}_i - \frac{1}{N}\sum_{i=1}^{N}x_{ki}\frac{1}{N}\sum_{i=1}^{N}\widehat{u}_i = \frac{1}{N}\sum_{i=1}^{N}x_{ki}\widehat{u}_i$$

si $\frac{1}{N}\sum_{i=1}^{N}\widehat{u}_i = 0.$

Notations vectorielles :

On notera xi : $\begin{pmatrix} 1 \\ x_{1i} \\ \vdots \\ x_{Ki} \end{pmatrix}$ | $\beta = \begin{pmatrix} a \\ b_1 \\ \vdots \\ b_K \end{pmatrix}$ sous formes de vecteurs

La transposée se note $\mathbf{x}_i^\mathsf{T} = (1, x_{1i}, \ldots, x_{Ki})$

Et le produit scalaire entre les deux devient notre équation :

$$\mathbf{x}_i \bullet \beta = \mathbf{x}_i^\mathsf{T}\beta = a + b_1 x_1 + b_2 x_2 + \ldots + b_K x_K$$

Les équations normales prennent une forme réduite :

$$\sum_{i=1}^{N} \mathbf{x}_i (y_i - \mathbf{x}_i^\mathsf{T} \widehat{\beta}) = 0$$

$$\sum_{i=1}^{N} \mathbf{x}_i y_i = \left(\sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^\mathsf{T} \right) \widehat{\beta}$$

Exemple vectoriel

$$
\begin{cases}
y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \ldots + \beta_k x_{k1} + u_1 \\
\ldots \\
y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + u_i \\
\ldots \\
y_N = \beta_0 + \beta_1 x_{1N} + \beta_2 x_{2N} + \ldots + \beta_k x_{kN} + u_N
\end{cases}
$$

Notations matricielles :

On appelle X la matrice de toutes les observation pour toute les variables:

$$X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{k1} & \cdots & x_{K1} \\ \vdots & & & \vdots & & \vdots \\ 1 & x_{1i} & \cdots & x_{ki} & \cdots & x_{Ki} \\ \vdots & & & \vdots & & \vdots \\ 1 & x_{1N} & \cdots & x_{kN} & \cdots & x_{KN} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1^\mathsf{T} \\ \vdots \\ \mathbf{x}_N^\mathsf{T} \end{pmatrix}$$

La première colonne contient des 1 pour l'intercept.

On note y le vecteur des variables dépendantes :

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = (y_1, \ldots, y_N)^\mathsf{T}.$$

$$
\begin{aligned}
X^{\mathsf{T}}\mathbf{y} &= \sum_{i=1}^{N} \mathbf{x}_i y_i \\
&= \sum_{i=1}^{N} \begin{pmatrix} 1 \\ x_{1i} \\ \vdots \\ x_{Ki} \end{pmatrix} y_i = \sum_{i=1}^{N} \begin{pmatrix} y_i \\ x_{1i} y_i \\ \vdots \\ x_{Ki} y_i \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
X^{\mathsf{T}}X &= (\mathbf{x}_1, \ldots, \mathbf{x}_N) \begin{pmatrix} \mathbf{x}_1^{\mathsf{T}} \\ \vdots \\ \mathbf{x}_N^{\mathsf{T}} \end{pmatrix} \\
&= \sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^{\mathsf{T}} \\
&= \begin{pmatrix}
N & \sum_{i=1}^{N} x_{1i} & \cdots & \sum_{i=1}^{N} x_{ki} & \cdots & \sum_{i=1}^{N} x_{Ki} \\
\vdots & & & \vdots & & \vdots \\
\sum_{i=1}^{N} x_{1i} & \sum_{i=1}^{N} x_{1i}^2 & \cdots & \sum_{i=1}^{N} x_{1i} x_{ki} & \cdots & \sum_{i=1}^{N} x_{1i} x_{Ki} \\
\vdots & & & \vdots & & \vdots \\
\sum_{i=1}^{N} x_{Ki} & \sum_{i=1}^{N} x_{Ki} x_{1i} & \cdots & \sum_{i=1}^{N} x_{Ki} x_{ki} & \cdots & \sum_{i=1}^{N} x_{Ki}^2
\end{pmatrix}
\end{aligned}
$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_i \\ \dots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1i} & x_{2i} & \dots & x_{ki} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1N} & x_{2N} & \dots & x_{kN} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \dots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_i \\ \dots \\ u_N \end{bmatrix}$$

$$[N\times 1] \qquad [N\times(k+1)] \qquad [(k+1)\times 1] \qquad [N\times 1]$$

$$\boxed{\underset{(N,1)}{Y} = \beta_0 + \beta_1 \underset{(N,1)}{X_1} + \cdots + \beta_K \underset{(N,1)}{X_K} + \underset{(N,1)}{u} = \underset{(N,K+1)}{X} \underset{(K+1,1)}{\beta} + \underset{(N,1)}{u}}$$

Exemple matriciel

- Soit, sous sa forme réduite : terminologie

$$\boxed{y} = \boxed{X}\boxed{\beta} + \boxed{u}$$

y	X	β	u
Variable dépendante Variable expliquée Variable réponse Variable régressée	Variables indépendantes Variables explicatives Variables de contrôle Régresseurs	Coefficients Paramètres Valeur de l'estimateur	Terme d'erreur Résidu
Observée *Variable certaine*	**Observées** *Variables certaines*	**À Estimer…** *Var. aléatoires*	**Inobservables** *Var. aléatoires*
Input (information disponible)		Output	Détermine le comportement des données

Matrice Singulière

- Une matrice carrée n*n est dite non singulière ou inversible si il existe une matrice inverse telle que AB =BA= In, ou I est la matrice identité (ne contient que des zéros et des 1 sur la diagonale).
- Si c'est le cas on peut déterminer *B* uniquement à partir de *A* et on l'appelle inverse de *A*, notée A^{-1}.
- Une Matrice carrée non inversible est dite singulière ou dégénérée, c'est le cas si son déterminant est zéro.
- La matrice A est inversible si toutes ses colonnes (ou lignes) sont indépendantes linéairement (ie on ne peut pas reconstruire une colonne en additionnant et multipliant les autres).

La condition de rang

- L'OLS est un estimateur solution a un système d'équations linéaires:

$$\sum_{i=1}^{N} \mathbf{x}_i y_i = \left(\sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^{\mathsf{T}} \right) \widehat{\beta}$$

$$X^{\mathsf{T}} \mathbf{y} = X^{\mathsf{T}} X \widehat{\beta}$$

- La solution existe et est unique si la matrice $X^{\mathsf{T}} X = \sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^{\mathsf{T}}$ est inversible, c'est la condition de rang.
- On peut démontrer que cela implique que les colonnes de X doivent être linéairement indépendantes (un régresseur ne doit pas) dépendre d'un autre.
- Si un répresseur dépend d'un autre on parle de multicolinéarité. Elle peut être parfaite (x2= 2*x1) ou non (dans ce cas la solution n'est pas fiable car on ne sait pas qui de X1 ou 2 influe vraiment sur y)

L'estimateur OLS

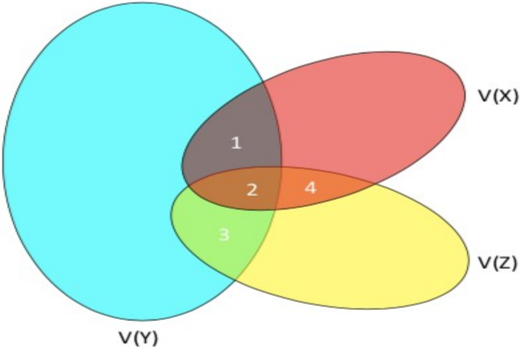
Si la condition de rang est vérifiée alors l'OLS admet pour unique solution l'estimateur :

$$\widehat{\beta} = \left(\sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^{\mathsf{T}}\right)^{-1} \sum_{i=1}^{N} \mathbf{x}_i y_i = (X^{\mathsf{T}} X)^{-1} X^{\mathsf{T}} \mathbf{y}$$

Hypothèses classiques

- Toutes les hypothèses du MLS
- + H10 : pas de multicolinéarité
  - Pas de relation linéaire parfaite entre les variables explicatives
  - Exemple : on ne peut pas avoir $x2 = 2x1$
- + H11 : la matrice X est de rang plein
  - Pour que (X'X) soit inversible
  - Pas d'observation avec données manquantes
    - sinon suppression des lignes ou imputation

Hypothèses classiques



Dans un modèle de régression multiple :

(1)+(2) = COV(Y,X)
(2)+(3) = COV(Y,Z)

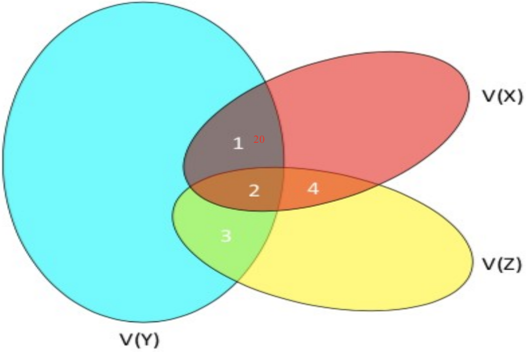
(1)+(2)+(3) / SCT = R^2

(1) = COV(Y,X |Z)

(3) = COV(Y,Z | X)

(2) + (4) = COV(X,Z)

Hypothèses classiques



Dans un modèle
de régression multiple :

(1)+(2) = COV(Y,X)
(2)+(3) = COV(Y,Z)

(1)+(2)+(3) / SCT = R^2

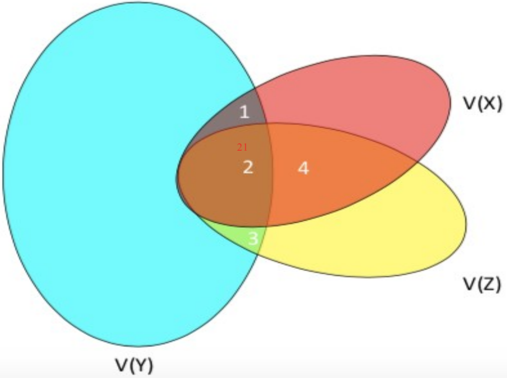
(1) = COV(Y,X | Z)

(3) = COV(Y,Z | X)

(2) + (4) = COV(X,Z)

(2) = **Multicolinéarité**

Hypothèses classiques



Dans un modèle
de régression multiple :

Si multicolinéarité **FORTE**

- Le R^2 diminue
- V(.) et σ des coeffs. β_X et β_Z diminuent ⇔ *Perte de significativité*
- X et Z apportent une information commune

Prediction

La prédiction de **y** est :

$$\begin{aligned}\widehat{\mathbf{y}} &= (\widehat{y}_1, \dots, \widehat{y}_N)^\top \\ &= X\widehat{\beta} = X(X^\top X)^{-1}X^\top \mathbf{y} = P_X \mathbf{y}\end{aligned}$$

On note $P_X = X(X^\top X)^{-1}X^\top \in \mathbb{R}^{N\times N}$ la projection orthogonale.

P_X a les propriétés suivantes:

$$\begin{aligned} P_X^\top &= P_X \quad \text{(symmetrie)}, \\ P_X P_X &= P_X \quad \text{(idempotence)} \end{aligned}$$

Le vecteur $\widehat{\mathbf{y}}$ est la projection orthogonale de **y** sur l'espace linéaire des combinaisons des colonnes de X (vecteur Xβ).

Résidus

Le vecteur des résidus :

$$\begin{aligned} \widehat{\mathbf{u}} &= (\widehat{u}_1, ..., \widehat{u}_N)^\top \\ &= \mathbf{y} - \widehat{\mathbf{y}} = \mathbf{y} - P_X\mathbf{y} = (I_N - P_X)\,\mathbf{y} \equiv M_X\mathbf{y}. \end{aligned}$$

La matrice $M_X = I_N - P_X \in \mathrm{R}^{N\times N}$ est un autre projecteur.

$$\begin{aligned} M_X &= I_N - P_X = M_X^\top \\ M_X M_X &= M_X \end{aligned}$$

À noter : $M_X X = 0_{N\times K} \in \mathrm{R}^{N\times K}$ et $M_X P_X = 0_{N\times N} \in \mathrm{R}^{N\times N}$.

Ce qui implique que le vecteur des résidus est orthogonal au vecteur des prédictions et des régresseurs :

$$\sum_{i=1}^{N} \widehat{y}_i \widehat{u}_i = \widehat{\mathbf{y}}^\top \widehat{\mathbf{u}} = (P_X\mathbf{y})^\top M_X\mathbf{y} = \mathbf{y}^\top P_X^\top M_X\mathbf{y} = \mathbf{y}^\top P_X M_X\mathbf{y} = 0,$$

Analysis of variance
(AN

$$\sum_{i=1}^{N}(y_i-\overline{y})^2=\sum_{i=1}^{N}(\widehat{y}_i-\overline{y})^2+\sum_{i=1}^{N}\widehat{u}_i^2$$

$SCT = SCE + SCR$.

variance totale = variance expliquée + variance inexpliquée.

Le coefficient de détermination

- A partir de l'équation d'analyse de la variance
- Permet d'estimer la qualité d'ajustement du modèle linéaire,

$$R2 = \frac{.*/}{.*0} = 1 - \frac{.*1}{.*0}$$

- Le R2 est
  - Toujours compris entre 0 et 1.
  - La part des variations de y expliquée par le modèle
- On peut comparer les R2 de deux modèles seulement si on a :
  - Les mêmes observations de y (sinon, on ne considère pas la même variance à expliquer)
  - Le même nombre de variables explicatives

Le coefficient de détermination

- Attention : le R2 augmente mécaniquement quand on ajoute une variable explicative.
- Solutions :
  - Compléter le R2 par une approche en termes de tests
    - cf. tests de Fisher
  - Calculer d'autres indicateurs de qualité prédictive :
    - Le R2 ajusté ($\overline{R}^2$)
    - Le Akaike's Information Criterion (AIC)
    - Schwartz's Information Criteria (SIC)

Le R2 ajusté

$$\overline{R}^2 = 1 - \frac{SSR/(N-K-1)}{SST/(N-1)}$$

- Plus le modèle minimise la variance du résidu plus la qualité de l'ajustement sera bonne.
- Si on augmente k : SCR va baisser mais (n-1-k) aussi...
- Pas de hausse systématique du R^2 avec k
- Il peut être négatif (si le fit est plus mauvais qu'une ligne horizontal, rare), il ne peut donc pas s'interpréter comme le R2
- Critère à maximiser

Les critères d'information : AIC

- Akaike's Information Criterion

$$AIC = e^{2(k+1)/n} \times \frac{SCR}{n}$$

  - Le modèle préféré sera le modèle avec le plus faible AIC
  - Le AIC pénalise encore plus l'ajout d'une variable supplémentaire que $R2$
  - Critère à minimiser

Les critères d'information : SIC

- Schwartz's Information Criteria

$$SIC = n^{(k+1)/n} \times \frac{SCR}{n}$$

  - Le modèle préféré sera le modèle avec le plus faible SIC.
  - Le SIC pénalise encore plus l'ajout d'une variable supplémentaire que $R2$ et le AIC.
  - Critère à minimiser

Exemple

- La fonction de production Cobb-Douglas

$$Y = A.K^{\alpha}.L^{\beta} \qquad avec \quad (\alpha+\beta=1)$$

- Question de recherche
  - Les rendements d'échelles sont-ils constants dans le secteur de l'industrie métallurgique aux EU ?
  - Données de 1957 pour N=25 Etats (Aigner, et al. 1977)
  - 3 variables (valeur ajoutée, capital, travail)
  - Comment estimer les paramètres du modèle ?

- La fonction de production Cobb-Douglas

$$Y = A.K^{\alpha}.L^{\beta} \quad avec \quad (\alpha+\beta=1)$$

Etape 1
Linéariser le modèle

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L$$

Exemple

- La fonction de production Cobb-Douglas

$$Y = A.K^{\alpha}.L^{\beta} \quad avec \quad (\alpha+\beta=1)$$

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L$$

$$\ln Y_i = \ln A + \alpha \ln K_i + \beta \ln L_i + u_i$$

Etape 2
Ecrire le modèle
Économétrique
pour l'échantillon

Exemple

- La fonction de production Cobb-Douglas

$$Y = A.K^{\alpha}.L^{\beta} \quad avec \quad (\alpha+\beta=1)$$

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L$$

$$\ln Y_i = \ln A + \alpha \ln K_i + \beta \ln L_i + u_i$$

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$$

Etape 3
Renommer les variable
pour retrouver
une forme propice
aux MCO

Exemple

- La fonction de production Cobb-Douglas

$$Y = A.K^{\alpha}.L^{\beta} \qquad avec \quad (\alpha+\beta=1)$$

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L$$

$$\ln Y_i = \ln A + \alpha \ln K_i + \beta \ln L_i + u_i$$

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$$

$$y = X\beta + u$$

Etape 4
Le modèle
Econométrique
sous forme matricielle

Exemple

	A	B	C	D
1	state	valueadded	capital	labor
2	Alabama	126,148	3,804	31,551
3	California	3201,486	185,446	452,844
4	Connecticut	690,670	39,712	124,074
5	Florida	56,296	6,547	19,181
6	Georgia	304,531	11,530	45,534
7	Illinois	723,028	58,987	88,391
8	Indiana	992,169	112,884	148,530
9	Iowa	35,796	2,698	8,017
10	Kansas	494,515	10,360	86,189
11	Kentucky	124,948	5,213	12,000
12	Louisiana	73,328	3,763	15,900
13	Maine	29,467	1,967	6,470
14	Maryland	415,262	17,546	69,342
15	Massachusetts	241,530	15,347	39,416
16	Michigan	4079,554	435,105	490,384
17	Missouri	652,085	32,840	84,831
18	NewJersey	667,113	33,292	83,033
19	NewYork	940,430	72,974	190,094
20	Ohio	1611,899	157,978	259,916
21	Pennsylvania	617,579	34,324	98,152
22	Texas	527,413	22,736	109,728
23	Virginia	174,394	7,173	31,301
24	Washington	636,948	30,807	87,963
25	WestVirginia	22,700	1,543	4,063
26	Wisconsin	349,711	22,001	52,818
27				

La première ligne contient le nom des variables

Il y a donc 26-1 =25 observations

Exemple

	A	B	C	D	E	F	G
1	state	valueadded	capital	labor	lny	lnk	lnl
2	Alabama	126,148	3,804	31,551	4,837	1,336	3,452
3	California	3201,486	185,446	452,844	8,071	5,223	6,116
4	Connecticut	690,670	39,712	124,074	6,538	3,682	4,821
5	Florida	56,296	6,547	19,181	4,031	1,879	2,954
6	Georgia	304,531	11,530	45,534	5,719	2,445	3,818
7	Illinois	723,028	58,987	88,391	6,583	4,077	4,482
8	Indiana	992,169	112,884	148,530	6,900	4,726	5,001
9	Iowa	35,796	2,698	8,017	3,578	0,993	2,082
10	Kansas	494,515	10,360	86,189	6,204	2,338	4,457
11	Kentucky	124,948	5,213	12,000	4,828	1,651	2,485
12	Louisiana	73,328	3,763	15,900	4,295	1,325	2,766
13	Maine	29,467	1,967	6,470	3,383	0,677	1,867
14	Maryland	415,262	17,546	69,342	6,029	2,865	4,239
15	Massachusetts	241,530	15,347	39,416	5,487	2,731	3,674
16	Michigan	4079,554	435,105	490,384	8,314	6,076	6,195
17	Missouri	652,085	32,840	84,831	6,480	3,492	4,441
18	NewJersey	667,113	33,292	83,033	6,503	3,505	4,419
19	NewYork	940,430	72,974	190,094	6,846	4,290	5,248
20	Ohio	1611,899	157,978	259,916	7,385	5,062	5,560
21	Pennsylvania	617,579	34,324	98,152	6,426	3,536	4,587
22	Texas	527,413	22,736	109,728	6,268	3,124	4,698
23	Virginia	174,394	7,173	31,301	5,161	1,970	3,444
24	Washington	636,948	30,807	87,963	6,457	3,428	4,477
25	WestVirginia	22,700	1,543	4,063	3,122	0,434	1,402
26	Wisconsin	349,711	22,001	52,818	5,857	3,091	3,967
27							

On créée de nouvelles variables :

lny = ln(valueadded)

lnk = ln(capital)

lnl = (ln(labor)

Exemple

	A	B	C	D	E	F	G
1	state	valueadded	capital	labor	lny	lnk	lnl
2	Alabama	126,148	3,804	31,551	4,837	1,336	3,452
3	California	3201,486	185,446	452,844	8,071	5,223	6,116
4	Connecticut	690,670	39,712	124,074	6,538	3,682	4,821
5	Florida	56,296	6,547	19,181	4,031	1,879	2,954
6	Georgia	304,531	11,530	45,534	5,719	2,445	3,818
7	Illinois	723,028	58,987	88,391	6,583	4,077	4,482
8	Indiana	992,169	112,884	148,530	6,900	4,726	5,001
9	Iowa	35,796	2,698	8,017	3,578	0,993	2,082
10	Kansas	494,515	10,360	86,189	6,204	2,338	4,457
11	Kentucky	124,948	5,213	12,000	4,828	1,651	2,485
12	Louisiana	73,328	3,763	15,900	4,295	1,325	2,766
13	Maine	29,467	1,967	6,470	3,383	0,677	1,867
14	Maryland	415,262	17,546	69,342	6,029	2,865	4,239
15	Massachusetts	241,530	15,347	39,416	5,487	2,731	3,674
16	Michigan	4079,554	435,105	490,384	8,314	6,076	6,195
17	Missouri	652,085	32,840	84,831	6,480	3,492	4,441
18	NewJersey	667,113	33,292	83,033	6,503	3,505	4,419
19	NewYork	940,430	72,974	190,094	6,846	4,290	5,248
20	Ohio	1611,899	157,978	259,916	7,385	5,062	5,560
21	Pennsylvania	617,579	34,324	98,152	6,426	3,536	4,587
22	Texas	527,413	22,736	109,728	6,268	3,124	4,698
23	Virginia	174,394	7,173	31,301	5,161	1,970	3,444
24	Washington	636,948	30,807	87,963	6,457	3,428	4,477
25	WestVirginia	22,700	1,543	4,063	3,122	0,434	1,402
26	Wisconsin	349,711	22,001	52,818	5,857	3,091	3,967
27							

Le vecteur y
contenant les y_i

Exemple

	A	B	C	D	E	F	G	H
1	state	valueadded	capital	labor	lny	lnk	lnl	const
2	Alabama	126,148	3,804	31,551	4,837	1,336	3,452	1
3	California	3201,486	185,446	452,844	8,071	5,223	6,116	1
4	Connecticut	690,670	39,712	124,074	6,538	3,682	4,821	1
5	Florida	56,296	6,547	19,181	4,031	1,879	2,954	1
6	Georgia	304,531	11,530	45,534	5,719	2,445	3,818	1
7	Illinois	723,028	58,987	88,391	6,583	4,077	4,482	1
8	Indiana	992,169	112,884	148,530	6,900	4,726	5,001	1
9	Iowa	35,796	2,698	8,017	3,578	0,993	2,082	1
10	Kansas	494,515	10,360	86,189	6,204	2,338	4,457	1
11	Kentucky	124,948	5,213	12,000	4,828	1,651	2,485	1
12	Louisiana	73,328	3,763	15,900	4,295	1,325	2,766	1
13	Maine	29,467	1,967	6,470	3,383	0,677	1,867	1
14	Maryland	415,262	17,546	69,342	6,029	2,865	4,239	1
15	Massachusetts	241,530	15,347	39,416	5,487	2,731	3,674	1
16	Michigan	4079,554	435,105	490,384	8,314	6,076	6,195	1
17	Missouri	652,085	32,840	84,831	6,480	3,492	4,441	1
18	NewJersey	667,113	33,292	83,033	6,503	3,505	4,419	1
19	NewYork	940,430	72,974	190,094	6,846	4,290	5,248	1
20	Ohio	1611,899	157,978	259,916	7,385	5,062	5,560	1
21	Pennsylvania	617,579	34,324	98,152	6,426	3,536	4,587	1
22	Texas	527,413	22,736	109,728	6,268	3,124	4,698	1
23	Virginia	174,394	7,173	31,301	5,161	1,970	3,444	1
24	Washington	636,948	30,807	87,963	6,457	3,428	4,477	1
25	WestVirginia	22,700	1,543	4,063	3,122	0,434	1,402	1
26	Wisconsin	349,711	22,001	52,818	5,857	3,091	3,967	1
27								

La matrice X
contenant les x_i
et la constante

Le vecteur y
contenant les y_i

Exemple

- Réduisons un peu le volume…

$$y = \begin{bmatrix} 4{,}837 \\ 8{,}071 \\ 6{,}653 \\ \dots \\ 5{,}857 \end{bmatrix} \quad , \quad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \dots & \dots & \dots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$

Exemple

$$X' = \begin{bmatrix} 1{,}336 & 5{,}223 & 3{,}682 & \dots & 3{,}091 \\ 3{,}452 & 6{,}116 & 4{,}821 & \dots & 3{,}967 \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \quad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \dots & \dots & \dots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$

Exemple

$$X' = \begin{bmatrix} 1,336 & 5,223 & 3,682 & \ldots & 3,091 \\ 3,452 & 6,116 & 4,821 & \ldots & 3,967 \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix} \quad X = \begin{bmatrix} 1,336 & 3,452 & 1 \\ 5,223 & 6,116 & 1 \\ 3,682 & 4,821 & 1 \\ \ldots & \ldots & \ldots \\ 3,091 & 3,967 & 1 \end{bmatrix}$$

$$X'X = \begin{bmatrix} \boxed{\quad} & \boxed{\quad} & \boxed{\quad} \\ \boxed{\quad} & \boxed{\quad} & \boxed{\quad} \\ \boxed{\quad} & \boxed{\quad} & \boxed{\quad} \end{bmatrix}$$

Rappel sur le calcul matriciel :
multiplication = somme des termes
en lignes x colonnes

Exemple

$$X' = \begin{bmatrix} 1{,}336 & 5{,}223 & 3{,}682 & \dots & 3{,}091 \\ 3{,}452 & 6{,}116 & 4{,}821 & \dots & 3{,}967 \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \quad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \dots & \dots & \dots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$

$$X'X = \begin{bmatrix} \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \end{bmatrix}$$

Rappel sur le calcul matriciel :
multiplication = somme des termes
en lignes x colonnes

Exemple

$$X' = \begin{bmatrix} 1{,}336 & 5{,}223 & 3{,}682 & \dots & 3{,}091 \\ 3{,}452 & 6{,}116 & 4{,}821 & \dots & 3{,}967 \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \qquad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \dots & \dots & \dots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$

$$X'X = \begin{bmatrix} \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \end{bmatrix}$$

Rappel sur le calcul matriciel : multiplication = somme des termes en lignes x colonnes

Exemple

$$X' = \begin{bmatrix} 1{,}336 & 5{,}223 & 3{,}682 & \dots & 3{,}091 \\ 3{,}452 & 6{,}116 & 4{,}821 & \dots & 3{,}967 \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \quad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \dots & \dots & \dots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$

$$X'X = \begin{bmatrix} \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \end{bmatrix}$$

Rappel sur le calcul matriciel :
multiplication = somme des termes
en lignes x colonnes

Exemple

$$X' = \begin{bmatrix} 1{,}336 & 5{,}223 & 3{,}682 & \dots & 3{,}091 \\ 3{,}452 & 6{,}116 & 4{,}821 & \dots & 3{,}967 \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \qquad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \dots & \dots & \dots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$

$$X'X = \begin{bmatrix} \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \end{bmatrix}$$

Rappel sur le calcul matriciel :
multiplication = somme des termes
en lignes x colonnes

Exemple

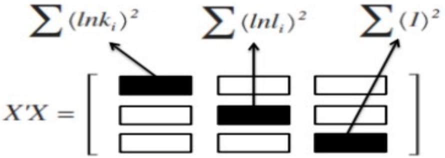
$$X' = \begin{bmatrix} 1{,}336 & 5{,}223 & 3{,}682 & \dots & 3{,}091 \\ 3{,}452 & 6{,}116 & 4{,}821 & \dots & 3{,}967 \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \qquad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \dots & \dots & \dots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$

$$X'X = \begin{bmatrix} \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \\ \boxed{} & \boxed{} & \boxed{} \end{bmatrix}$$

Rappel sur le calcul matriciel :
multiplication = somme des termes
en lignes x colonnes

Exemple

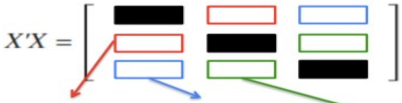
$$X' = \begin{bmatrix} 1{,}336 & 5{,}223 & 3{,}682 & \ldots & 3{,}091 \\ 3{,}452 & 6{,}116 & 4{,}821 & \ldots & 3{,}967 \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix} \quad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \ldots & \ldots & \ldots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$



Rappel sur le calcul matriciel :
multiplication :
- Sur la diagonale sont les termes carrés

Exemple

$$X' = \begin{bmatrix} 1{,}336 & 5{,}223 & 3{,}682 & \dots & 3{,}091 \\ 3{,}452 & 6{,}116 & 4{,}821 & \dots & 3{,}967 \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \quad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \dots & \dots & \dots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$



Rappel sur le calcul matriciel :
multiplication :
- Sur la diagonale sont les termes carrés
- Ailleurs, sont les termes croisés

$\sum(lnk)\times(lnl)$ $\quad \sum(lnk)\times(1)$ $\quad \sum(lnl)\times(1)$

Exemple

$$X' = \begin{bmatrix} 1{,}336 & 5{,}223 & 3{,}682 & \dots & 3{,}091 \\ 3{,}452 & 6{,}116 & 4{,}821 & \dots & 3{,}967 \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \quad X = \begin{bmatrix} 1{,}336 & 3{,}452 & 1 \\ 5{,}223 & 6{,}116 & 1 \\ 3{,}682 & 4{,}821 & 1 \\ \dots & \dots & \dots \\ 3{,}091 & 3{,}967 & 1 \end{bmatrix}$$

$$X'X = \begin{bmatrix} 271{,}2 & 339{,}9 & 73{,}9 \\ 339{,}9 & 442{,}7 & 100{,}6 \\ 73{,}9 & 100{,}6 & 25 \end{bmatrix}$$

X'X est une matrice symétrique

Exemple

$$(X'X)^{-1} = \begin{bmatrix} \mathbf{0,206} & & \\ -0,231 & \mathbf{0,287} & \\ 0,324 & -0,472 & \mathbf{0,982} \end{bmatrix}$$

$(X'X)^{-1}$ est aussi symétrique,
Pour simplifier, on ne présente
que les valeurs sous la diagonale

Voir par ailleurs comment
inverser une matrice

Exemple

$$(X'X)^{-1} = \begin{bmatrix} 0,206 & & \\ -0,231 & 0,287 & \\ 0,324 & -0,472 & 0,982 \end{bmatrix} \qquad X'y = \begin{bmatrix} 476,7 \\ 625,5 \\ 145,3 \end{bmatrix} \begin{array}{l} \to \sum (\ln y)\times(\ln k) \\ \to \sum (\ln y)\times(\ln l) \\ \to \sum (\ln y)\times(1) \end{array}$$

Exemple

$$(X'X)^{-1} = \begin{bmatrix} \mathbf{0{,}206} & & \\ -0{,}231 & \mathbf{0{,}287} & \\ 0{,}324 & -0{,}472 & \mathbf{0{,}982} \end{bmatrix} \quad \mathbf{X'y} = \begin{bmatrix} 476{,}7 \\ 625{,}5 \\ 145{,}3 \end{bmatrix}$$

$$\hat{\beta} = (X'X)^{-1}(X'y) = \begin{bmatrix} 0{,}206 & -0{,}231 & 0{,}324 \\ -0{,}231 & 0{,}287 & -0{,}472 \\ 0{,}324 & -0{,}472 & 0{,}982 \end{bmatrix} \begin{bmatrix} 476{,}7 \\ 625{,}5 \\ 145{,}3 \end{bmatrix}$$

Exemple

$$(X'X)^{-1} = \begin{bmatrix} \mathbf{0,206} & & \\ -0,231 & \mathbf{0,287} & \\ 0,324 & -0,472 & \mathbf{0,982} \end{bmatrix} \quad \mathbf{X'y} = \begin{bmatrix} 476,7 \\ 625,5 \\ 145,3 \end{bmatrix}$$

$$\hat{\beta} = (X'X)^{-1}\,(X'y) = \begin{bmatrix} 0,245 \\ 0,805 \\ 1,844 \end{bmatrix}$$

Exemple

$$(X'X)^{-1} = \begin{bmatrix} \mathbf{0,206} & & \\ -0,231 & \mathbf{0,287} & \\ 0,324 & -0,472 & \mathbf{0,982} \end{bmatrix} \quad X'y = \begin{bmatrix} 476,7 \\ 625,5 \\ 145,3 \end{bmatrix}$$

$$\hat{\beta} = (X'X)^{-1}(X'y) = \begin{bmatrix} 0,245 \\ 0,805 \\ 1,844 \end{bmatrix} \begin{matrix} \rightarrow \text{Coefficient de } \ln k \\ \rightarrow \text{Coefficient de } \ln l \\ \rightarrow \text{Valeur de la constante} \end{matrix}$$

Exemple

$$\ln y_i = 1{,}844 + 0{,}245 \ln x_{1i} + 0{,}805 \ln x_{2i}$$

$$Y_i = (e^{1{,}844}).K_i^{0{,}245} L_i^{0{,}805}$$

- Les rendements sont-ils constants ?
  - $(\hat{\alpha}+\hat{\beta}) = (0{,}245 + 0{,}805) = 1{,}045 \approx 1$
  - OUI : La somme des facteurs d'échelle est proche de 1,
- Impact des facteurs de production sur la plus-value ...
- Interprétation des coefficients ?
  - Quand le capital augmente de 10%, le niveau de production augmente de 2,5%, tcepa
  - Quand le travail augmente de 10%, le niveau de production augmente de 8%, tcepa
  - Sans capital ni travail le log du niveau de la production est de 1,8 ....?