Introduction à l'économétrie

Applications en économie et en finance

Jaouad Madkour
jmadkour@uae.ac.ma

Faculté des sciences juridiques, économiques et sociales de Tanger
Département des sciences économiques et de gestion

Site Web : jmadkour.com
Facebook : www.facebook.com/jmadkour
YouTube : www.youtube.com/JaouadMadkour

25 janvier 2024

كلية العلوم القانونية والاقتصادية
والاجتماعية – طنجة
FSJES TANGER

Introduction

- Qu'est ce que l'économétrie ?
- Économétrie *vs* Data Science
- Démarche de l'économétrie
- Notion de régression
- Références bibliographiques
- Plan du cours

Introduction

Qu'est ce que l'économétrie ?

Introduction

Qu'est ce que l'économétrie ?

Samuelson, Koopmans et Stone (1954), *Report of the Evaluative Committee for Econometrica*, Econometrica, volume 22, numéro 2, pp. 141-146

> *"L'économétrie peut être définie comme l'analyse quantitative des phénomènes économiques basée sur le développement concurrent de la théorie et de l'observation reliées par des méthodes appropriées de déduction."*

Tintner (1968), *Methodology of Mathematical Economics and Econometrics*, UCP, p. 74

> *"L'économétrie [...] consiste à appliquer les mathématiques statistiques aux données économiques pour fournir une base empirique aux modèles construits par l'économie mathématique et obtenir des résultats mesurés."*

Introduction

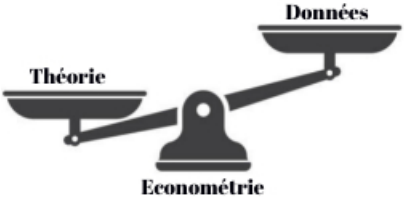
Qu'est ce que l'économétrie ?

Étymologie

Le terme *économétrie* est la fusion du terme **économie** et du terme **métrie** (= **mesure**).
Au sens littéral, l'économétrie signifie **mesure de l'économie**.

Définition

L'**économétrie** est l'application de méthodes statistiques à des données économiques pour **valider** ou **rejeter** des théories économiques.



Économétrie *vs* Data Science

Introduction

Économétrie *vs* Data Science

Joshua Angrist, Prix Nobel d'économie 2021

"L'économétrie est la science des données originelle. Avant qu'il y ait la science des données, il y avait l'économétrie."

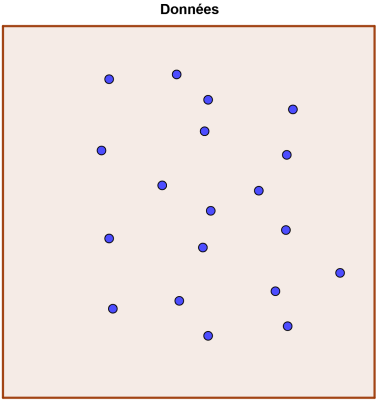
"Les entreprises cherchent des personnes pour faire des statistiques, mais les questions qui les intéressent sont des questions causales : quelles seront les conséquences d'un changement de prix ou d'un changement de stratégie marketing ? La meilleure formation pour cela est l'économétrie."

"Les data scientistes sont concernés par des méthodes d'ajustement pour faire des prévisions. Ainsi, tout modèle qui s'ajuste bien aux données passées pourrait être extrapolé vers l'avenir.
Les économètres s'intéressent plutôt à des relations causales. Le passé n'est pas nécessairement un bon guide car la variation d'une variable peut être associée à d'autres variables dont le changement peut potentiellement affecter les résultats."

Introduction

Économétrie *vs* Data Science

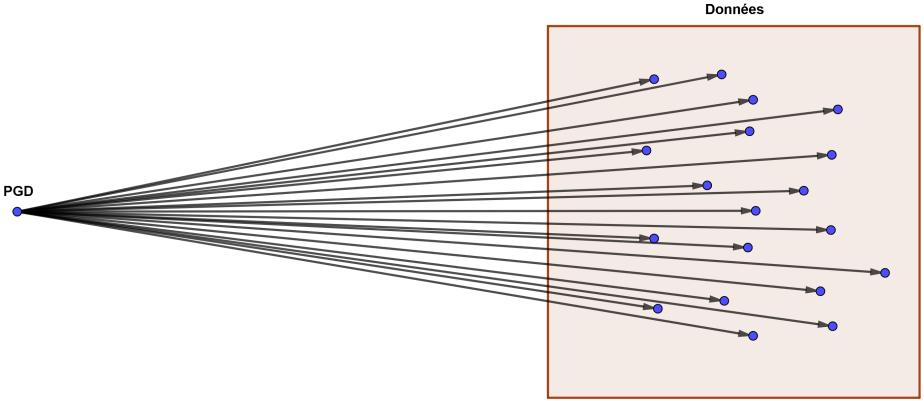
En Data Science, l'accent est mis sur les données tandis qu'en économétrie l'accent est mis le Processus Générateur des Données (PGD).



Introduction

Économétrie *vs* Data Science

En Data Science, l'accent est mis sur les données tandis qu'en économétrie l'accent est mis le Processus Générateur des Données (PGD).



Démarche de l'économétrie

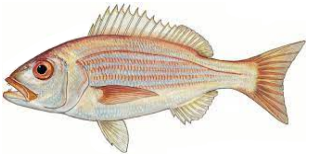
Introduction

Démarche de l'économétrie

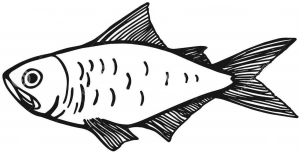
- Le **PGD** n'étant pas connu, l'économètre va l'approcher par un (ou des) modèle(s) ;
- Un **modèle** est une représentation simplifiée de la **réalité** ;
- L'économètre fait de la **modélisation économétrique** pour **expliquer** ou **prévoir** un phénomène.

> *"All models are wrong, but some are useful"*
>
> George Box (Statisticien britannique)

Réalité



Modèle 1



Modèle 2

Introduction

Démarche de l'économétrie

Méthodologie traditionnelle de l'économétrie :

1. Énoncé de la théorie ou des hypothèses ;
2. Spécification du modèle mathématique de la théorie ;
3. Spécification du modèle économétrique ;
4. Obtention des données ;
5. Estimation des paramètres du modèle économétrique ;
6. Test des hypothèses ;
7. Utilisation du modèle à des fins de prévision ou de contrôle.

Introduction

Démarche de l'économétrie

1. Énoncé de la théorie ou des hypothèses

Dans sa *Théorie générale de l'emploi, de l'intérêt et de la monnaie*, J.M. Keynes (économiste) annonce sa *loi psychologique fondamentale* comme suit :

> *"En moyenne et la plupart du temps, les hommes tendent à accroître leur consommation à mesure que leur revenu croît, mais non d'une quantité aussi grande que l'accroissement du revenu."*

Keynes suppose clairement une relation croissante entre la consommation C et le revenu R :

$$C = f(R) \qquad\qquad f'(R) > 0$$

Keynes suppose que la consommation C croît moins vite que le revenu R :

$$\Delta C < \Delta R$$

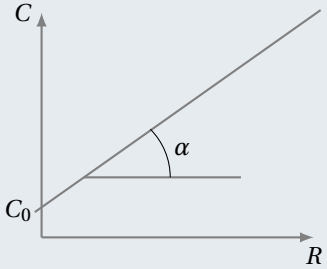
Keynes ne donne pas de formulation mathématique à la fonction f.

2. Spécification du modèle mathématique de la théorie

J. Hicks (économiste mathématicien) a formalisé mathématiquement la *Théorie générale* de Keynes. La fonction f prend une forme linéaire :

$$C = C_0 + c.R \text{ avec } 0 < c < 1 \tag{1}$$

- C : Consommation des ménages
- R : Revenu disponible
- C_0 : Consommation incompressible $(R = 0)$
- c : Propension marginale à consommer $(= dC/dR)$
- $c = \tan(\alpha) : 0 < c < 1 \Leftrightarrow 0 < \alpha < 45°$



3. Spécification du modèle économétrique

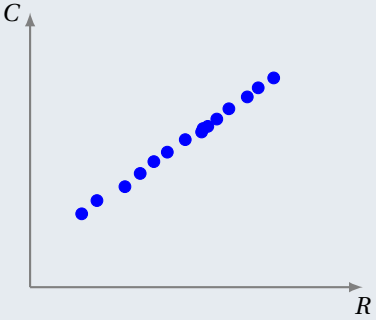
L'expression (1) suppose une relation déterministe entre la consommation C et le revenu R. Mais rien ne prouve, d'une part, que la consommation dépende exclusivement du revenu, car elle peut aussi dépendre de la taille des ménages et de l'âge des individus par exemple. Et rien ne garantit, d'autre part, que les variables économiques soient mesurées avec précision. Pour cela, un terme d'erreurs ε doit être inséré dans le modèle mathématique (1) pour tenir compte de ces deux problèmes. On obtient le modèle économétrique suivant :

$$C = C_0 + c.R + \varepsilon \tag{2}$$

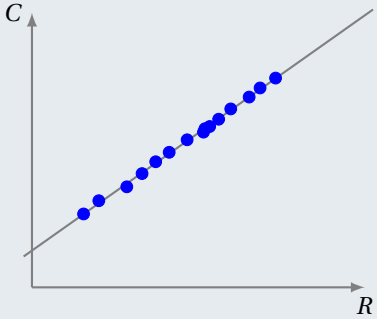
4. Obtention des données

Consommation privée (C) et PIB (R) aux USA entre 1982 et 1996 en milliards de dollars :

Année	C	R	Année	C	R
1982	3081.5	4620.3	1990	4132.2	6136.3
1983	3240.6	4803.7	1991	4105.8	6079.4
1984	3407.6	5140.1	1992	4219.8	6244.4
1985	3566.5	5323.5	1993	4343.6	6389.6
1986	3708.7	5487.7	1994	4486.0	6610.7
1987	3822.3	5649.5	1995	4595.3	6742.1
1988	3972.7	5865.2	1996	4714.1	6928.4
1989	4064.6	6062.0			



5. Estimation des paramètres du modèle économétrique



$$\hat{C} = -184.08 + 0.7064R \qquad (3)$$

Sur la période étudiée (1982-1996) :

- L'estimation du niveau moyen de la consommation est $\hat{C} = -184.08 + 0.7064R$
- L'estimation de la propension marginale à consommer est de $\hat{c} = 0.7064$
- L'estimation de la consommation incompressible est $\hat{C}_0 = -184.08$ milliards de dollars

L'estimation du modèle économétrique fournit un contenu empirique à la théorie économique.

6. Test des hypothèses

1. On suppose que l'expression (3) est une bonne approximation de la réalité ;
2. La théorie keynésienne prévoit une propension marginale à consommer comprise entre 0 et 1 ;
3. L'estimation économétrique de la propension marginale à consommer est de $\hat{c} = 0.7064$;
4. L'estimation économétrique confirme-t-elle la théorie keynésienne de la consommation ?

Afin de valider ou rejeter la théorie keynésienne, il faut tester deux hypothèses :

- L'estimation $\hat{c} = 0.7064$ est-elle significativement différente de 0 ?
- L'estimation $\hat{c} = 0.7064$ est-elle significativement inférieure à 1 ?

7. Utilisation du modèle à des fins de prévision ou de contrôle

On suppose que les tests précédents **n'ont pas rejeté** la théorie keynésienne. L'expression (3) peut être utilisée pour prédire les valeurs futures de la consommation C sur la base des valeurs futures du revenu disponible R.

Exemple :

On suppose que l'on veuille prédire la dépense de consommation moyenne en 1997, le PIB étant de 7269.8 milliards de dollars. Il suffit de remplacer R par 7269.8 dans l'expression (3), on obtient :

$$\begin{aligned}\hat{C}_{1997} &= -184.08 + 0.7064 \times 7269.8 \\ &= 4951.31 \text{ milliards de dollars}\end{aligned}$$

Notion de régression

Introduction

Notion de régression

Régression

L'objet principal de l'économétrie est l'**analyse de régression** :

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \cdots + \beta_k \mathbf{x}_k + \varepsilon$$

- $\mathbf{y}$: la **régressande** (ou encore **variable dépendante**, **expliquée** ou **endogène**)
- $\mathbf{x}$: les **régresseurs** (ou encore **variables indépendantes**, **explicatives** ou **exogènes**)
- β : les **paramètres** du modèle de régression
- ε : le **terme d'erreurs** (ou encore **innovations**, **perturbations** ou **chocs**)

Exemple :

Le modèle de consommation keynésien est **un modèle de régression linéaire simple** :

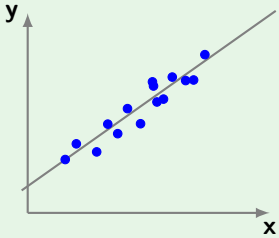
$$C = C_0 + cY + \varepsilon$$

Introduction

Notion de régression

Toute la question est de savoir...

- ... comment estimer les paramètres β ? et comment les interpréter ?
- ... comment tester la significativité (individuelle et jointe) des régresseurs $\mathbf{x}$?
- ... comment évaluer la qualité d'estimation du modèle ? et comment l'améliorer ?
- ... comment choisir entre des modèles concurrents ? et comment les utiliser ?

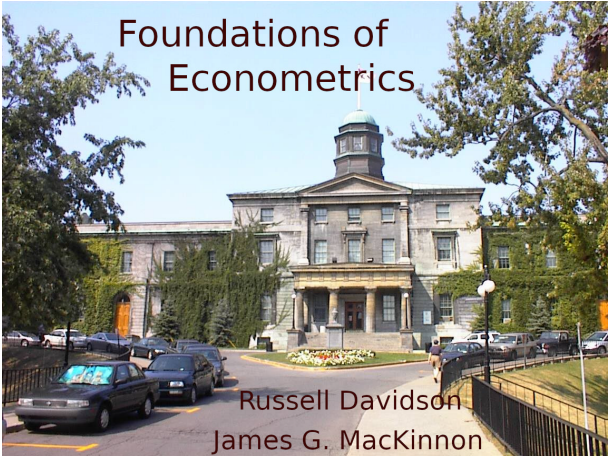


$$\mathbf{y} = a + b\mathbf{x} + \varepsilon$$

Références bibliographiques

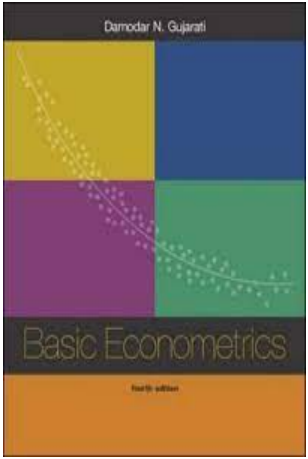
Introduction

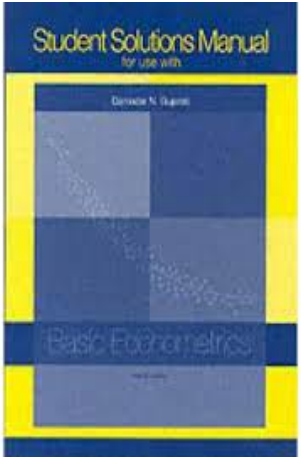
Références bibliographiques

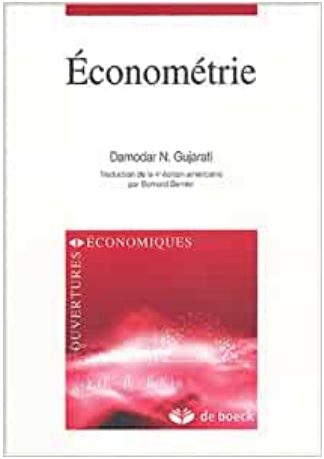


Introduction

Références bibliographiques







Introduction

Références bibliographiques

- Davidson, R. and MacKinnon, J.G. (2021). Foundations of Econometrics.
- Gujarati, D.N. (2004). Basic Econometrics. Fourth edition. The McGraw-Hill Companies ;
- Gujarati, D.N. (2004). Student Solutions Manual for Use with Basic Econometrics (Fourth edition). McGraw-Hill higher education ;
- Gujarati, D.N. and Bernier, B. (2004). Économétrie. Ouvertures économiques. De Boeck Supérieur.

Plan du cours

Introduction

Plan du cours

Chapitre 1 : Les fondements statistiques de l'économétrie

Chapitre 2 : Les fondements algébriques de l'économétrie

Chapitre 3 : La géométrie des Moindres Carrés Ordinaires

Chapitre 4 : Les propriétés statistiques des Moindres Carrés Ordinaires

Chapitre 5 : Les tests d'hypothèses dans les modèles de régression linéaire

Chapitre 6 : Les intervalles de confiance

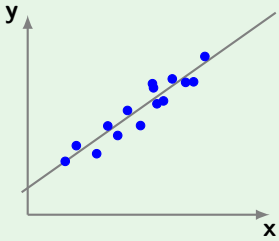
Chapitre 1 :

Les fondements statistiques de l'économétrie

1. Distributions statistiques à une dimension
2. Distributions statistiques à deux dimensions

Les fondements statistiques de l'économétrie

Objectif :



$$\mathbf{y} = a + b\mathbf{x} + \varepsilon \Longrightarrow \hat{\mathbf{y}} = \hat{a} + \hat{b}\mathbf{x}$$

$$\hat{a} = \overline{\mathbf{y}} - \hat{b}\overline{\mathbf{x}} \qquad \hat{b} = \frac{cov(\mathbf{x}, \mathbf{y})}{V(\mathbf{x})}$$

Section 1 :

Distributions statistiques à une dimension

1. Centre d'une distribution statistique
2. Dispersion d'une distribution statistique
3. Asymétrie d'une distribution statistique
4. Aplatissement d'une distribution statistique

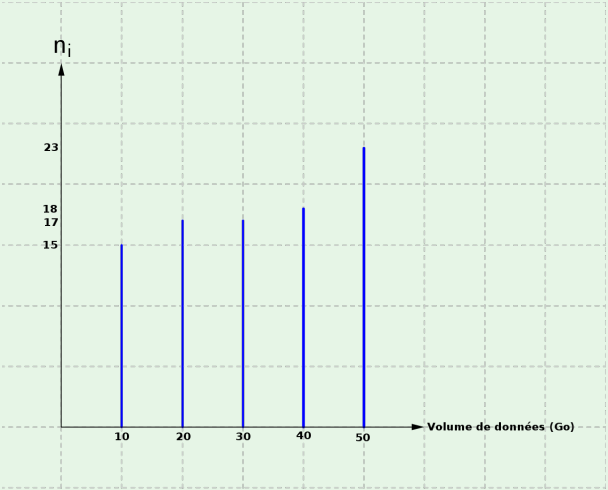
Exemple : Besoin mensuel en volumes de données (en Go) d'un échantillon d'étudiants

10	40	30	50	10	40	20	30	20	10
20	50	10	40	30	10	30	40	30	10
40	50	30	20	40	50	10	40	30	50
10	50	20	50	30	50	30	20	40	10
30	50	50	10	40	40	50	40	20	50
30	20	40	10	50	50	50	20	50	30
40	20	50	20	50	30	50	40	30	40
10	20	10	20	30	30	40	40	20	20
20	30	10	50	20	50	50	40	50	10

Go	10	20	30	40	50
n_i	15	17	17	18	23

Les fondements statistiques de l'économétrie

Distributions statistiques à une dimension : Centre d'une distribution statistique



Sous-section 1 : Centre d'une distribution statistique

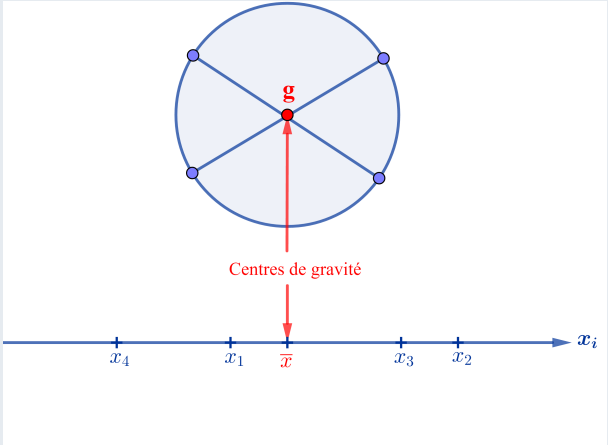
Moyenne arithmétique

Moyenne arithmétique

La *moyenne arithmétique* d'une variable statistique, notée $\overline{x}$, est la valeur qui concentre toute l'information contenue dans la distribution statistique. C'est l'équivalent d'un centre de gravité en géométrie. Elle est donnée par la formule suivante :

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i = \sum_{i=1}^{k} f_i x_i$$

Centres de gravité :



Exemple : Besoin mensuel en volumes de données (en Go) d'un échantillon d'étudiants

x_i	n_i	$n_i x_i$
10	15	150
20	17	340
30	17	510
40	18	720
50	23	1150
Σ	90	2870

$$\begin{aligned}\overline{x} &= \frac{1}{n}\sum_{i=1}^{5} n_i x_i \\ &= \frac{2870}{90} \\ &\approx 31.89 \text{ Go}\end{aligned}$$

Les étudiants questionnés ont besoin, en moyenne, de 31.89 Go de données par mois.

Propriétés :

Soient X et Y deux variables statistiques quantitatives prenant respectivement les valeurs $x_1, x_2, \cdots, x_k$ et $y_1, y_2, \cdots, y_k$ avec $k \in \mathbb{N}^*$ et a et b deux constantes réelles.

La moyenne arithmétique vérifie les propriétés suivantes :

$$\overline{ax + by} = a\overline{x} + b\overline{y}$$

$$\sum_{i=1}^{k} n_i (x_i - \overline{x}) = 0$$

Sous-section 2 : Dispersion d'une distribution statistique

Écart absolu moyen

Variance

Écart-type

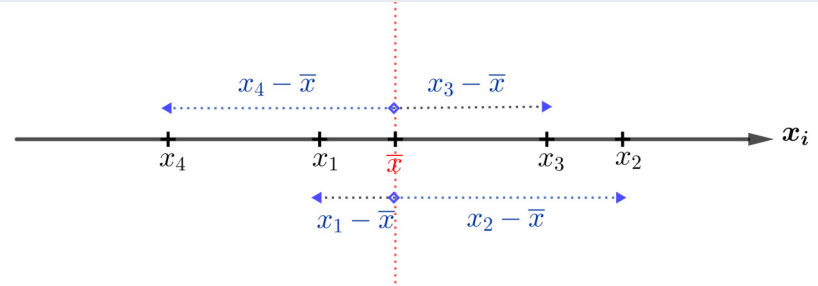
Coefficient de variation

Exemple introductif : Taux de rentabilité d'actifs financiers

Actif A					Actif B					Actif C				
0%	0%	0%	0%	0%	-2%	-1%	0%	1%	2%	-4%	-3%	0%	3%	4%
		$\overline{x} = 0\%$					$\overline{x} = 0\%$					$\overline{x} = 0\%$		

La moyenne arithmétique ne donne aucune information sur la *dispersion* des données !

Écarts par rapport à la moyenne arithmétique :



Écart moyen

L'*écart moyen* d'une variable statistique, notée EM, est la moyenne arithmétique des écarts des observations de la variable X par rapport au centre généralement mesuré par la moyenne arithmétique $\overline{x}$:

$$EM = \frac{1}{n}\sum_{i=1}^{k} n_i(x_i - \overline{x}) = 0$$

Remarque

L'inconvénient majeur de l'écart moyen est qu'il est toujours nul, quel que soit la dispersion de la variable statistique étudiée. Cela est dû au fait que les écarts $x_i - \overline{x}$ positifs et les écarts $x_i - \overline{x}$ négatifs se compensent. La somme des écarts est en conséquence nulle. Il faudrait donc trouver un moyen de neutraliser l'effet des signes négatifs pour éviter cette compensation de signes. La première solution qui s'impose est d'appliquer une valeur absolue à ces écarts et considérer des écarts absolus $|x_i - \overline{x}|$ à la place des écarts $x_i - \overline{x}$.

Écart absolu moyen

L'*écart absolu moyen* d'une variable statistique, notée EAM, est la moyenne arithmétique des valeurs absolues des écarts des observations de la variable X par rapport au centre généralement mesuré par la moyenne arithmétique $\overline{x}$:

$$\text{EAM} = \frac{1}{n}\sum_{i=1}^{k} n_i \left|x_i - \overline{x}\right| \geq 0$$

Remarque

L'inconvénient de l'écart absolu moyen est qu'il n'est pas dérivable. Il faudrait donc trouver une autre alternative qui neutralise l'effet des signes négatifs des écarts $x_i - \overline{x}$ et qui soit en même temps dérivable. La puissance 2 est une bonne alternative.

Exemple : Besoin mensuel en volumes de données (en Go) d'un échantillon d'étudiants

x_i	n_i	$\lvert x_i - \overline{x} \rvert$	$n_i \lvert x_i - \overline{x} \rvert$
10	15	21.89	328.35
20	17	11.89	202.13
30	17	1.89	32.13
40	18	8.11	145.98
50	23	18.11	416.53
Σ	90	–	1125.12

$$\text{EAM} = \frac{1}{n}\sum_{i=1}^{5} n_i \left| x_i - \overline{x} \right| = \frac{1125.12}{90} \approx 12.50 \text{ Go}$$

Les besoins mensuels des étudiants en volumes de donnée s'éloignent de la moyenne $\overline{x} = 31.89$ Go de 12.50 Go en moyenne.

Variance & écart-type

La *variance* d'une variable statistique X, notée σ_x^2 ou $V(X)$, est la moyenne des carrés des écarts à la moyenne $\overline{x}$:

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{k} n_i \left(x_i - \overline{x}\right)^2$$

L'*écart-type* d'une variable statistique X, noté σ_x, est la racine carrée de la variance : $\sigma_x = \sqrt{\sigma_x^2}$.

Théorème de König-Huygens : Formule réduite de la variance

La variance peut se réécrire $\sigma_x^2 = \overline{x^2} - \overline{x}^2$ avec

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2$$

Démonstration :

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{k} n_i\left(x_i - \overline{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{k} n_i\left(x_i^2 - 2x_i\overline{x} + \overline{x}^2\right)$$

$$= \frac{1}{n}\sum_{i=1}^{k}\left(n_i x_i^2 - n_i 2x_i\overline{x} + n_i\overline{x}^2\right) = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - \frac{1}{n}\sum_{i=1}^{k} n_i 2x_i\overline{x} + \frac{1}{n}\sum_{i=1}^{k} n_i\overline{x}^2$$

$$= \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - 2\overline{x}\left(\frac{1}{n}\sum_{i=1}^{k} n_i x_i\right) + \overline{x}^2\left(\frac{1}{n}\sum_{i=1}^{k} n_i\right) = \overline{x^2} - 2\overline{x}.\overline{x} + \overline{x}^2 = \overline{x^2} - \overline{x}^2$$

Exemple : Besoin mensuel en volumes de données (en Go) d'un échantillon d'étudiants

x_i	n_i	$n_i x_i^2$	$n_i\left(x_i-\overline{x}\right)^2$
10	15	1500	7187.5815
20	17	6800	2403.3257
30	17	15300	60.7257
40	18	28800	1183.8978
50	23	57500	7543.3583
Σ	90	109900	18378.889

$$\sigma_x^2=\frac{1}{n}\sum_{i=1}^{5} n_i\left(x_i-\overline{x}\right)^2=\frac{18378.889}{90}\approx 204.21\Rightarrow\sigma=\sqrt{\sigma_x^2}\approx 14.29 \text{ Go}$$

$$\sigma_x^2=\overline{x^2}-\overline{x}^2=\frac{109900}{90}-(31.89)^2\approx 204.14\Rightarrow\sigma=\sqrt{\sigma_x^2}\approx 14.29 \text{ Go}$$

Les besoins mensuels des étudiants en volumes de donnée s'éloignent de la moyenne $\overline{x}=31.89$ Go de 14.29 Go en moyenne.

Remarque :

La variance n'a pas de signification concrète, il s'agit simplement d'une étape intermédiaire dans le calcul de l'écart-type. Néanmoins, elle peut être utilisée comme mesure de dispersion tout en prenant garde à bien interpréter l'unité de mesure transformée à la puissance 2 car ceci peut induire en erreur. A titre d'exemples, le mètre qui est une unité de mesure des longueurs se transforme en mètre carré qui est une unité de mesure des surfaces et le litre se transforme en litre carré qui n'a pas de signification concrète.

Propriétés :

Soit X une variable statistique quantitative et a et b deux constantes réelles. La variance vérifie les propriétés suivantes :

$$V(aX) = a^2 V(X)$$

$$V(X+b) = V(X)$$

$$V(aX+b) = a^2 V(X)$$

Démonstrations :

$$V(aX) = \frac{1}{n}\sum_{i=1}^{k} n_i\left(ax_i - \overline{ax}\right)^2 = \frac{1}{n}\sum_{i=1}^{k} n_i\left(ax_i - a\overline{x}\right)^2$$

$$= \frac{1}{n}\sum_{i=1}^{k} n_i\left[a(x_i - \overline{x})\right]^2 = \frac{1}{n}\sum_{i=1}^{k} n_i a^2\left(x_i - \overline{x}\right)^2$$

$$= a^2\frac{1}{n}\sum_{i=1}^{k} n_i\left(x_i - \overline{x}\right)^2 = a^2 V(X)$$

X	aX	n_i
x_1	ax_1	n_1
x_2	ax_2	n_2
$\vdots$	$\vdots$	$\vdots$
x_i	ax_i	n_i
$\vdots$	$\vdots$	$\vdots$
x_k	ax_k	n_k
	Σ	n

Démonstrations (suite) :

$$
\begin{aligned}
V(X+b) &= \frac{1}{n}\sum_{i=1}^{k} n_i \left(x_i + b - \overline{x+b}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{k} n_i \left[x_i + b - (\overline{x}+b)\right]^2 \\
&= \frac{1}{n}\sum_{i=1}^{k} n_i \left(x_i + b - \overline{x} - b\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{k} n_i \left(x_i - \overline{x}\right)^2 = V(X)
\end{aligned}
$$

X	$X+b$	n_i
x_1	x_1+b	n_1
x_2	x_2+b	n_2
$\vdots$	$\vdots$	$\vdots$
x_i	x_i+b	n_i
$\vdots$	$\vdots$	$\vdots$
x_k	x_k+b	n_k
	Σ	n

Démonstrations (suite) :

$$
\begin{aligned}
V(aX+b) &= V(aX) \\
&= a^2 V(X)
\end{aligned}
$$

X	$aX+b$	n_i
x_1	ax_1+b	n_1
x_2	ax_2+b	n_2
$\vdots$	$\vdots$	$\vdots$
x_i	ax_i+b	n_i
$\vdots$	$\vdots$	$\vdots$
x_k	ax_k+b	n_k
	Σ	n

Exemples :

$$V(2X) = 4V(X)$$

$$V(X+3) = V(X)$$

$$V(2X+3) = 4V(X)$$

Remarque

L'inconvénient majeur de l'écart type σ est qu'il dépend de l'unité de mesure de la variable statistique étudiée. De ce fait, il ne permet pas la comparaison des dispersions de variables quantitatives qui ne sont pas exprimées dans les mêmes unités de mesure.

Coefficient de variation

Le coefficient de variation, noté CV, est une mesure de dispersion relative, il est égal au rapport entre l'écart-type et la moyenne arithmétique, c'est un nombre sans dimension :

$$CV = \frac{\sigma}{\overline{x}}$$

Le coefficient de variation permet de comparer les dispersions de variables quantitatives qui ne sont pas exprimées dans les mêmes unités de mesure

Exemples : Besoin mensuel des étudiants en volumes de données et en durées d'appels

$\overline{x} = 31.39$ Go

$\sigma_x = 14.29$ Go

$$CV_x = \frac{\sigma}{\overline{x}} = \frac{14.29\ \cancel{Go}}{31.39\ \cancel{Go}} \approx 0.4552$$

$\overline{y} = 7.91$ h

$\sigma_y = 4.35$ h

$$CV_y = \frac{\sigma}{\overline{y}} = \frac{4.35\ \cancel{h}}{7.91\ \cancel{h}} \approx 0.5499$$

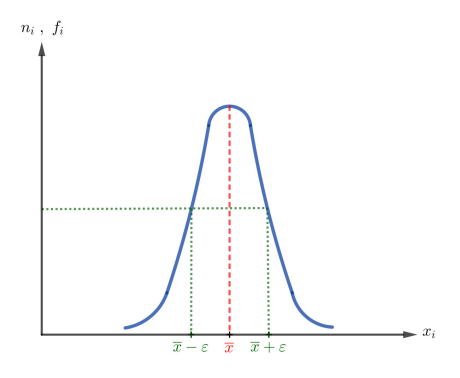
La distribution statistique des besoins mensuels des étudiants en durées d'appels est plus dispersée relativement à celle de leurs besoins mensuels en volumes de données.

Sous-section 3 : Asymétrie d'une distribution statistique

Coefficient de Skewness

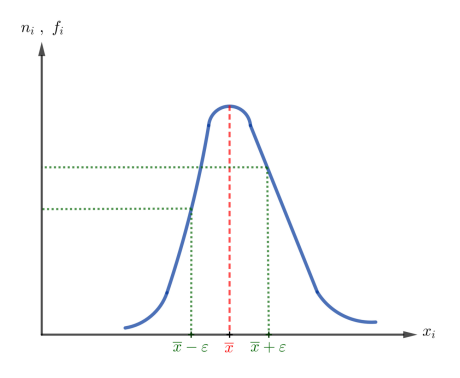
Distribution symétrique :

Les valeurs symétriques par rapport au centre de la distribution statistique ont les mêmes effectifs et les mêmes fréquences.



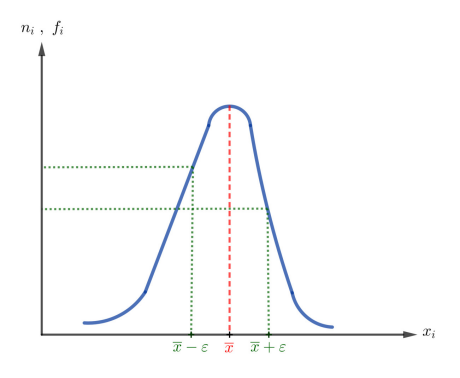
Distribution asymétrique à droite :

Les valeurs supérieures au centre de la distribution statistique sont plus fréquentes relativement à leurs symétriques.



Distribution asymétrique à gauche :

Les valeurs inférieures au centre de la distribution statistique sont plus fréquentes relativement à leurs symétriques.



Coefficient de *skewness* de Fisher :

$$\gamma_1 = \frac{1}{n}\sum_{i=1}^{k} n_i \left(\frac{x_i - \overline{x}}{\sigma}\right)^3$$

Interprétation :

- $\gamma_1 = 0$: la distribution est *symétrique* ;
- $\gamma_1 > 0$: la distribution est *asymétrique à droite* ;
- $\gamma_1 < 0$: la distribution est *asymétrique à gauche*.

Exemple : Besoin mensuel en volumes de données (en Go) d'un échantillon d'étudiants

x_i	n_i	$n_i\left(\frac{x_i-\overline{x}}{\sigma_x}\right)^3$
10	15	-33.75
20	17	-8.704
30	17	-0.017
40	18	3.888
50	23	50.531
Σ	90	11.948

$$\gamma_1 = \frac{1}{n}\sum_{i=1}^{5} n_i\left(\frac{x_i-\overline{x}}{\sigma_x}\right)^3 = \frac{11.948}{90} \approx 0.133$$

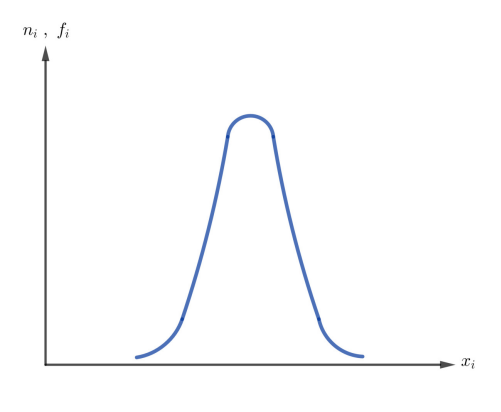
La distribution des besoins mensuels des étudiants en volumes de données est asymétrique à droite. Cela signifie que les étudiants qui ont des besoins mensuels en volumes de données supérieurs à la moyenne sont plus fréquents que ceux qui ont des besoins inférieurs à la moyenne.

Sous-section 4 : Aplatissement d'une distribution statistique

Coefficient de Kurtosis

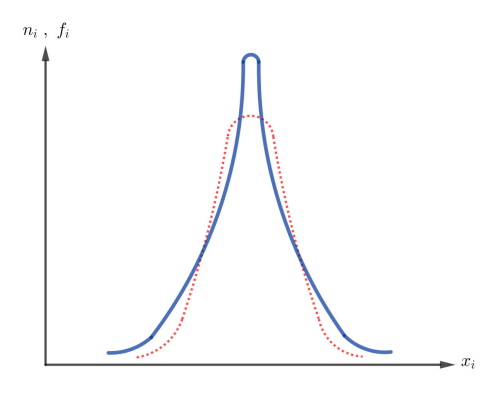
Distribution mésokurtique :

Les valeurs extrêmes et les valeurs centrales ont des fréquences modérées.



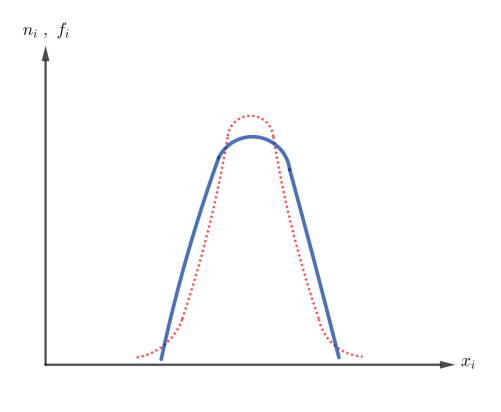
Distribution leptokurtique :

Les valeurs extrêmes et les valeurs centrales sont plus fréquentes relativement à une distribution normale (en pointillés).



Distribution platykurtique :

Les valeurs extrêmes et les valeurs centrales sont moins fréquentes relativement à une distribution normale (en pointillés).



Coefficient de *kurtosis* de Pearson :

$$\beta_2 = \frac{1}{n}\sum_{i=1}^{k} n_i \left(\frac{x_i - \overline{x}}{\sigma}\right)^4$$

Interprétation :

- $\beta_2 = 3$: distribution *mésokurtique* (courbe normale) ;
- $\beta_2 > 3$: distribution *leptokurtique* (moins aplatie que la normale) ;
- $\beta_2 < 3$: distribution *platykurtique* (plus aplatie que la normale).

Exemple : Besoin mensuel en volumes de données (en Go) d'un échantillon d'étudiants

x_i	n_i	$n_i\left(\frac{x_i-\overline{x}}{\sigma_x}\right)^4$
10	15	50.625
20	17	6.96
30	17	0.0017
40	18	2.333
50	23	65.69
Σ	90	125.6097

$$\beta_2 = \frac{1}{n}\sum_{i=1}^{5} n_i\left(\frac{x_i-\overline{x}}{\sigma_x}\right)^4 = \frac{125.6097}{90} \approx 1.40$$

La distribution des besoins mensuels des étudiants en volumes de données est platykurtique. Cela signifie que les valeurs extrêmes sont moins fréquentes relativement à une distribution normale.

Section 2 :

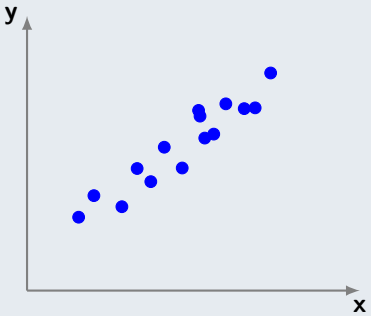
Distributions statistiques à deux dimensions

1. Nuage de points et tableau de contingence
2. Distributions marginales
3. Distributions conditionnelles
4. Distribution jointe
5. Ajustement linéaire d'un nuage de points
6. La méthode des moindres carrés ordinaires

Sous-section 1 :

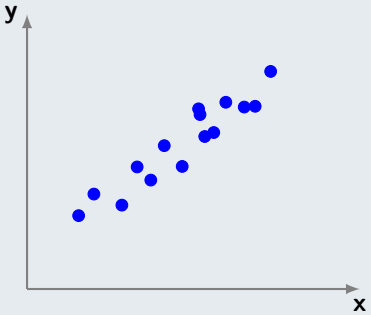
Nuage de points et tableau de contingence

Nuage de points *vs* tableau de contingence



x \ y	y_1	y_2	$\cdots$	y_j	$\cdots$	y_J	Σ
x_1	n_{11}	n_{12}	$\cdots$	n_{1j}	$\cdots$	n_{1J}	$n_{1.}$
x_2	n_{21}	n_{22}	$\cdots$	n_{2j}	$\cdots$	n_{2J}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	n_{i1}	n_{i2}	$\cdots$	n_{ij}	$\cdots$	n_{iJ}	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_I	n_{I1}	n_{I2}	$\cdots$	n_{Ij}	$\cdots$	n_{IJ}	$n_{I.}$
Σ	$n_{.1}$	$n_{.2}$	$\cdots$	$n_{.j}$	$\cdots$	$n_{.J}$	$n_{..}$

Nuage de points *vs* tableau de contingence



x	y	n_i
x_1	y_1	1
x_2	y_2	1
$\vdots$	$\vdots$	$\vdots$
x_i	y_i	1
$\vdots$	$\vdots$	$\vdots$
x_n	y_n	1
Σ		n

Sous-section 2 :

Distributions marginales

Distribution marginale de **x** :

x \ y	y_1	y_2	$\cdots$	y_j	$\cdots$	y_J	Σ
x_1	n_{11}	n_{12}	$\cdots$	n_{1j}	$\cdots$	n_{1J}	$n_{1.}$
x_2	n_{21}	n_{22}	$\cdots$	n_{2j}	$\cdots$	n_{2J}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	n_{i1}	n_{i2}	$\cdots$	n_{ij}	$\cdots$	n_{iJ}	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_I	n_{I1}	n_{I2}	$\cdots$	n_{Ij}	$\cdots$	n_{IJ}	$n_{I.}$
Σ	$n_{.1}$	$n_{.2}$	$\cdots$	$n_{.j}$	$\cdots$	$n_{.J}$	$n_{..}$

$$\overline{\mathbf{x}} = \frac{1}{n_{..}} \sum_{i=1}^{I} n_{i.} x_i \qquad V(\mathbf{x}) = \frac{1}{n_{..}} \sum_{i=1}^{I} n_{i.} \left(x_i - \overline{\mathbf{x}}\right)^2 = \overline{\mathbf{x}^2} - \overline{\mathbf{x}}^2$$

Distribution marginale de **y** :

x \ y	y_1	y_2	$\cdots$	y_j	$\cdots$	y_J	Σ
x_1	n_{11}	n_{12}	$\cdots$	n_{1j}	$\cdots$	n_{1J}	$n_{1.}$
x_2	n_{21}	n_{22}	$\cdots$	n_{2j}	$\cdots$	n_{2J}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	n_{i1}	n_{i2}	$\cdots$	n_{ij}	$\cdots$	n_{iJ}	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_I	n_{I1}	n_{I2}	$\cdots$	n_{Ij}	$\cdots$	n_{IJ}	$n_{I.}$
Σ	$n_{.1}$	$n_{.2}$	$\cdots$	$n_{.j}$	$\cdots$	$n_{.J}$	$n_{..}$

$$\overline{\mathbf{y}} = \frac{1}{n_{..}} \sum_{j=1}^{J} n_{.j} y_j \qquad V(\mathbf{y}) = \frac{1}{n_{..}} \sum_{j=1}^{J} n_{.j} \left(y_j - \overline{\mathbf{y}}\right)^2 = \overline{\mathbf{y}^2} - \overline{\mathbf{y}}^2$$

Sous-section 3 :

Distributions conditionnelles

Distributions conditionnelles de $\mathbf{x}|\mathbf{y} = y_j$, $j = 1, 2, \cdots, J$:

$\mathbf{x}$ \ $\mathbf{y}$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_J	Σ
x_1	n_{11}	n_{12}	$\cdots$	n_{1j}	$\cdots$	n_{1J}	$n_{1.}$
x_2	n_{21}	n_{22}	$\cdots$	n_{2j}	$\cdots$	n_{2J}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	n_{i1}	n_{i2}	$\cdots$	n_{ij}	$\cdots$	n_{iJ}	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_I	n_{I1}	n_{I2}	$\cdots$	n_{Ij}	$\cdots$	n_{IJ}	$n_{I.}$
Σ	$n_{.1}$	$n_{.2}$	$\cdots$	$n_{.j}$	$\cdots$	$n_{.J}$	$n_{..}$

$$\overline{\mathbf{x}}_j = \frac{1}{n_{.j}} \sum_{i=1}^{I} n_{ij} x_i \qquad V_j(\mathbf{x}) = \frac{1}{n_{..}} \sum_{i=1}^{I} n_{i.} \left(x_i - \overline{\mathbf{x}}_j\right)^2 = \overline{\mathbf{x}_j^2} - \overline{\mathbf{x}}_j^2$$

Distributions conditionnelles de $\mathbf{y}|\mathbf{x} = x_i$, $i = 1, 2, \cdots, I$:

$\mathbf{x}$ \ $\mathbf{y}$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_J	Σ
x_1	n_{11}	n_{12}	$\cdots$	n_{1j}	$\cdots$	n_{1J}	$n_{1.}$
x_2	n_{21}	n_{22}	$\cdots$	n_{2j}	$\cdots$	n_{2J}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	n_{i1}	n_{i2}	$\cdots$	n_{ij}	$\cdots$	n_{iJ}	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_I	n_{I1}	n_{I2}	$\cdots$	n_{Ij}	$\cdots$	n_{IJ}	$n_{I.}$
Σ	$n_{.1}$	$n_{.2}$	$\cdots$	$n_{.j}$	$\cdots$	$n_{.J}$	$n_{..}$

$$\overline{\mathbf{y}}_i = \frac{1}{n_{i.}} \sum_{j=1}^{J} n_{ij} y_j \qquad V_i(\mathbf{y}) = \frac{1}{n_{..}} \sum_{j=1}^{J} n_{.j} \left(y_j - \overline{\mathbf{y}}_i\right)^2 = \overline{\mathbf{y}_i^2} - \overline{\mathbf{y}}_i^2$$

Sous-section 4 :

Distribution jointe

Distribution jointe du couple $(\mathbf{x}, \mathbf{y})$

$\mathbf{x}$ \ $\mathbf{y}$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_J	Σ
x_1	n_{11}	n_{12}	$\cdots$	n_{1j}	$\cdots$	n_{1J}	$n_{1.}$
x_2	n_{21}	n_{22}	$\cdots$	n_{2j}	$\cdots$	n_{2J}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	n_{i1}	n_{i2}	$\cdots$	n_{ij}	$\cdots$	n_{iJ}	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_I	n_{I1}	n_{I2}	$\cdots$	n_{Ij}	$\cdots$	n_{IJ}	$n_{I.}$
Σ	$n_{.1}$	$n_{.2}$	$\cdots$	$n_{.j}$	$\cdots$	$n_{.J}$	$n_{..}$

$$\overline{\mathbf{xy}} = \frac{1}{n_{..}} \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} x_i y_j \qquad cov(\mathbf{x}, \mathbf{y}) = \frac{1}{n_{..}} \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \left(x_i - \overline{\mathbf{x}}\right)\left(y_j - \overline{\mathbf{y}}\right) = \overline{\mathbf{xy}} - \overline{\mathbf{x}}.\overline{\mathbf{y}}$$

Origine de la covariance $cov(\mathbf{x},\mathbf{y})$

La variance $V(\mathbf{x})$ mesure la dispersion d'une variable statistique $\mathbf{x}$ autour de sa moyenne $\overline{\mathbf{x}}$:

$$V(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{I} n_i \left(x_i - \overline{\mathbf{x}}\right)^2 = \overline{\mathbf{x}^2} - \overline{\mathbf{x}}^2$$

$$= \frac{1}{n}\sum_{i=1}^{I} n_i \left(x_i - \overline{\mathbf{x}}\right)\left(x_i - \overline{\mathbf{x}}\right) = \overline{\mathbf{x}.\mathbf{x}} - \overline{\mathbf{x}}.\overline{\mathbf{x}}$$

En substituant l'un des $\mathbf{x}$ par $\mathbf{y}$, on obtient une mesure de variation d'une variable $\mathbf{x}$ par rapport à une variable $\mathbf{y}$. Il s'agit de la covariance entre $\mathbf{x}$ et $\mathbf{y}$:

$$cov(\mathbf{x},\mathbf{y}) = \frac{1}{n}\sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij} \left(x_i - \overline{\mathbf{x}}\right)\left(y_j - \overline{\mathbf{y}}\right) = \overline{\mathbf{x}.\mathbf{y}} - \overline{\mathbf{x}}.\overline{\mathbf{y}}$$

Remarque

Contrairement à la variance $V(\mathbf{x})$ qui est par construction non négative, la covariance $cov(\mathbf{x},\mathbf{y})$ peut être positive, négative ou nulle :

- $cov(\mathbf{x},\mathbf{y}) > 0 \Leftrightarrow \mathbf{x}$ et $\mathbf{y}$ varient dans le même sens ;
- $cov(\mathbf{x},\mathbf{y}) < 0 \Leftrightarrow \mathbf{x}$ et $\mathbf{y}$ varient en sens inverses ;
- $cov(\mathbf{x},\mathbf{y}) = 0 \Leftrightarrow \mathbf{x}$ et $\mathbf{y}$ sont indépendantes l'une de l'autre.

Remarque

L'inconvénient majeur de la covariance $cov(\mathbf{x},\mathbf{y})$ est qu'elle dépend à la fois de l'unité de mesure de la variable $\mathbf{x}$ et de l'unité de mesure de la variable $\mathbf{y}$. Il en résulte que l'unité de mesure de la covariance $cov(\mathbf{x},\mathbf{y})$ est la multiplication de deux unités de mesure et qu'elle est de ce fait démunie de sens.

Coefficient de corrélation linéaire entre $\mathbf{x}$ et $\mathbf{y}$

Afin de neutraliser sa dépendance vis-à-vis de l'unité de mesure de $\mathbf{x}$ et de $\mathbf{y}$, on divise la covariance $cov(\mathbf{x},\mathbf{y})$ par l'écart-type de $\mathbf{x}$ et par l'écart-type de $\mathbf{y}$. On obtient **le coefficient de corrélation linéaire de Pearson** :

$$\rho(\mathbf{x},\mathbf{y}) = \frac{cov(\mathbf{x},\mathbf{y})}{\sqrt{V(\mathbf{x})}\sqrt{V(\mathbf{y})}}$$

Interprétation

Le coefficient de corrélation linéaire $\rho(\mathbf{x},\mathbf{y})$ est un nombre sans dimension qui mesure l'intensité de la relation linéaire entre deux variables quantitatives continues **x** et **y**. Il est compris entre -1 et 1 :

- $\rho(\mathbf{x},\mathbf{y}) = -1$: il y a une relation décroissante parfaitement linéaire entre **x** et **y** ;
- $\rho(\mathbf{x},\mathbf{y}) \to -1$: il y a une forte relation linéaire décroissante entre **x** et **y** ;
- $\rho(\mathbf{x},\mathbf{y}) \to 0^-$: il y a une faible relation linéaire décroissante entre **x** et **y** ;
- $\rho(\mathbf{x},\mathbf{y}) = 0$: il n'y a une aucune relation linéaire entre **x** et **y** ;
- $\rho(\mathbf{x},\mathbf{y}) \to 0^+$: il y a une faible relation linéaire croissante entre **x** et **y** ;
- $\rho(\mathbf{x},\mathbf{y}) \to 1$: il y a une forte relation linéaire croissante entre **x** et **y** ;
- $\rho(\mathbf{x},\mathbf{y}) = 1$: il y a une relation croissante parfaitement linéaire entre **x** et **y**.

Illustration graphique : $\rho = -1$

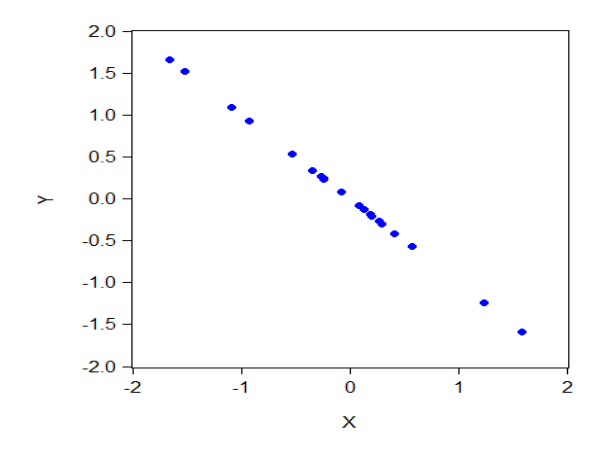


Illustration graphique : $\rho \to -1$

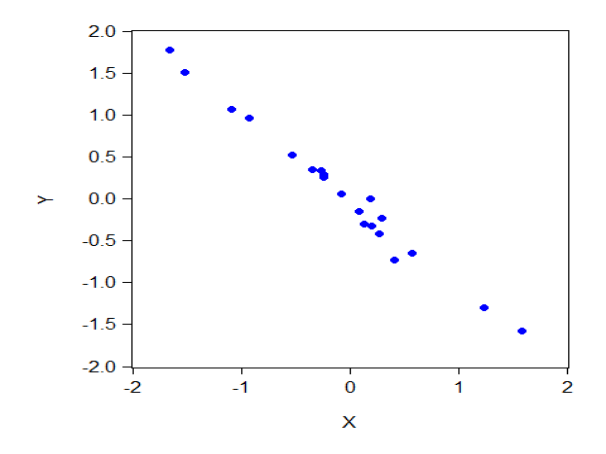


Illustration graphique : $\rho \to 0^-$

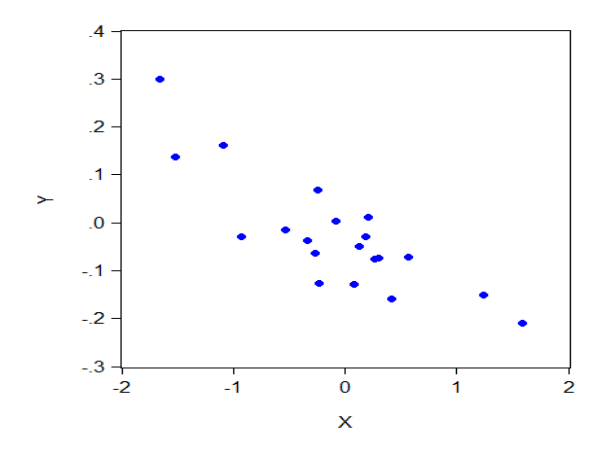


Illustration graphique : $\rho = 0$

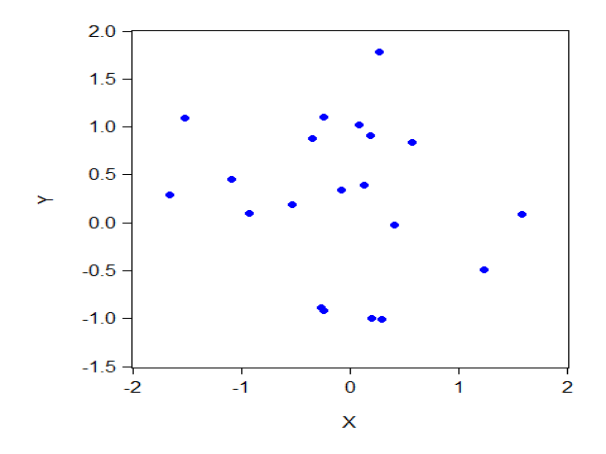


Illustration graphique : $\rho \to 0^+$

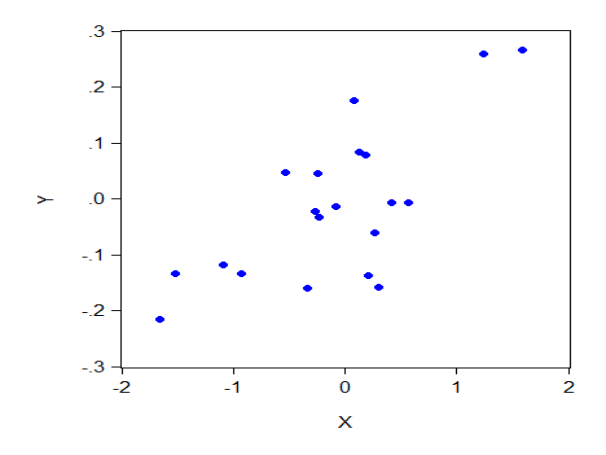


Illustration graphique : $\rho \to 1$

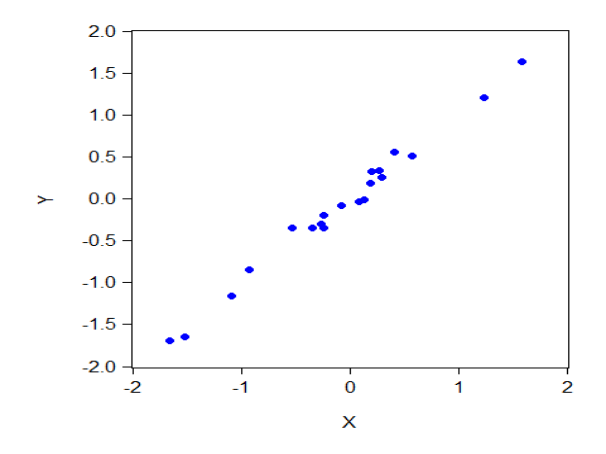
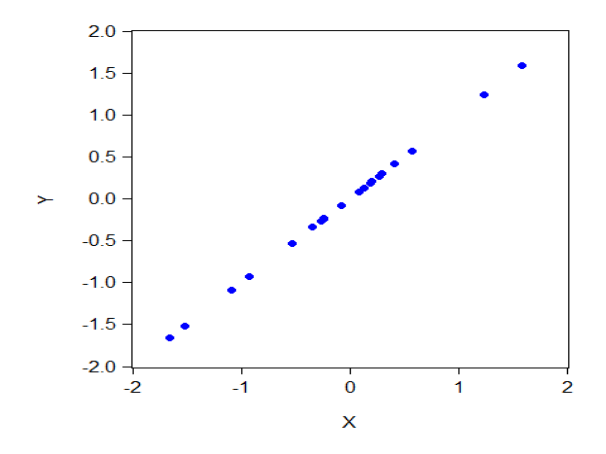


Illustration graphique : $\rho = 1$



Sous-section 5 :

Ajustement linéaire d'un nuage de points

Illustration graphique : $\rho = -1$

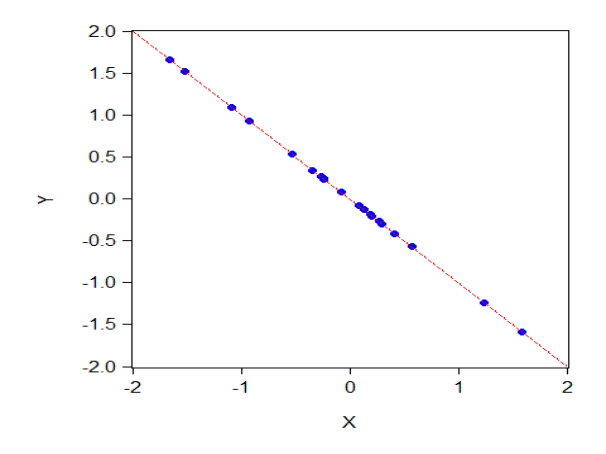


Illustration graphique : $\rho \to -1$

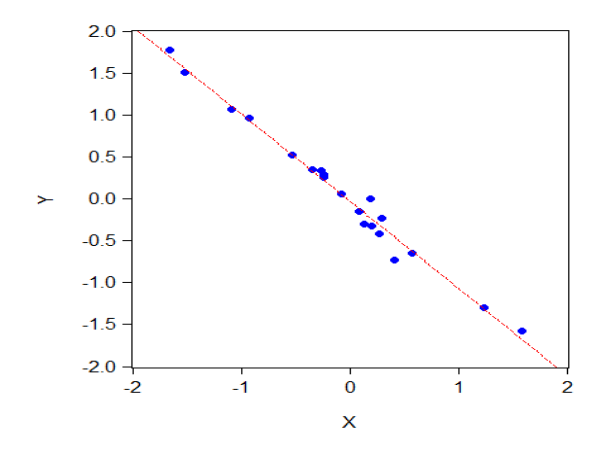


Illustration graphique : $\rho \to 0^-$

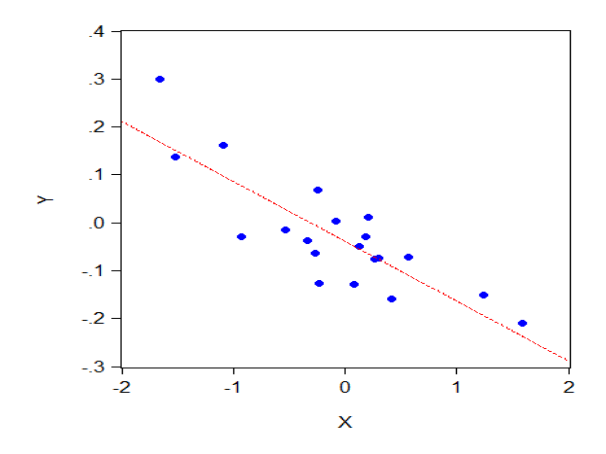


Illustration graphique : $\rho = 0$

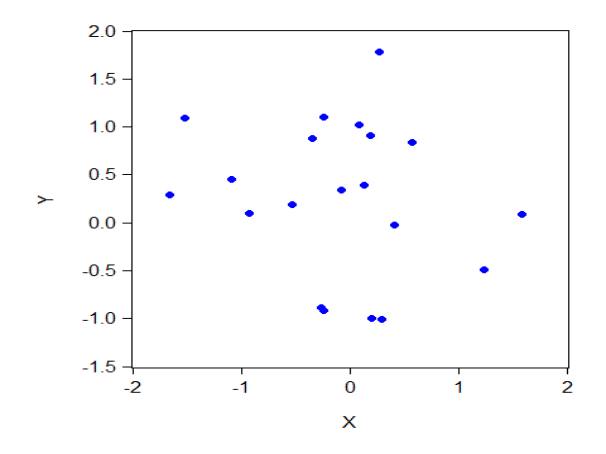


Illustration graphique : $\rho \to 0^+$

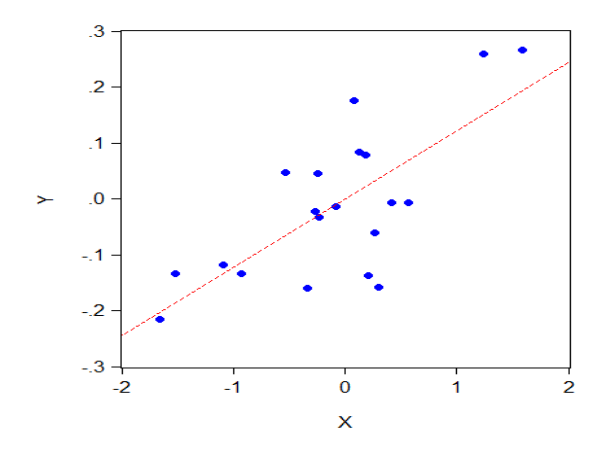


Illustration graphique : $\rho \to 1$

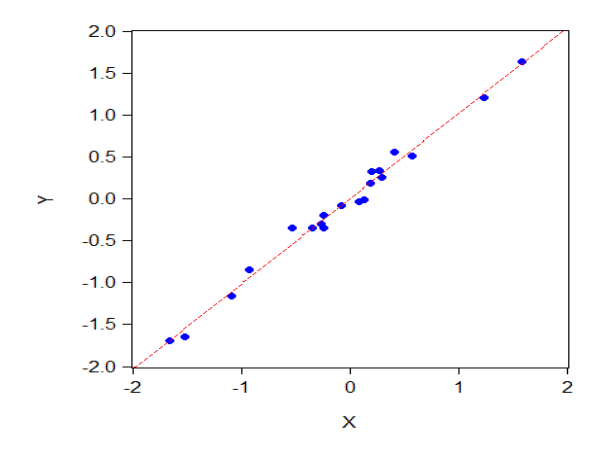
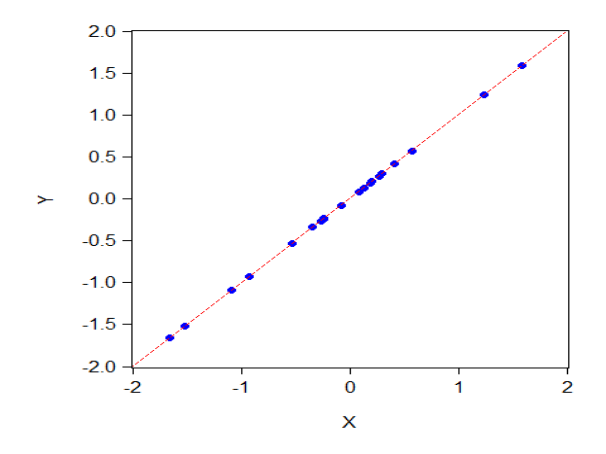


Illustration graphique : $\rho = 1$



Sous-section 6 :

La méthode des moindres carrés ordinaires (MCO)

Régression linéaire simple :

$$\mathbf{y} = a + b\mathbf{x} + \varepsilon$$

Régression linéaire simple pour l'observation i :

$$y_i = a + bx_i + \varepsilon_i \ , \ i = 1, 2, \cdots, n$$

Erreurs :

$$\varepsilon_i = y_i - a - bx_i \ , \ i = 1, 2, \cdots, n$$

Carrés des erreurs :

$$\varepsilon_i^2 = \left(y_i - a - bx_i\right)^2 \ , \ i = 1, 2, \cdots, n$$

Somme des carrés des erreurs :

$$f(a,b) \equiv \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left(y_i - a - bx_i\right)^2$$

Estimateurs des moindres carrés ordinaires :

$$\hat{a} = \arg\min_{a} f(a,b)$$

$$\hat{b} = \arg\min_{b} f(a,b)$$

Minimisation de la somme des carrés des erreurs $f(a,b)$ par rapport au paramètre a :

$$\begin{aligned}\frac{\partial f(a,b)}{\partial a} &= \frac{\partial}{\partial a}\sum_{i=1}^{n}\left(y_i - a - bx_i\right)^2 \\ &= \sum_{i=1}^{n}\frac{\partial}{\partial a}\left(y_i - a - bx_i\right)^2 \\ &= \sum_{i=1}^{n}(-2)\left(y_i - a - bx_i\right) \\ &= -2\sum_{i=1}^{n}\left(y_i - a - bx_i\right)\end{aligned}$$

$$\begin{aligned}&\left.\frac{\partial f(a,b)}{\partial a}\right|_{a=\hat{a}} = 0 \\ &\sum_{i=1}^{n}\left(y_i - \hat{a} - bx_i\right) = 0 \\ &\sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\hat{a} - \sum_{i=1}^{n} bx_i = 0 \\ &\frac{1}{n}\sum_{i=1}^{n} y_i - \frac{1}{n}\sum_{i=1}^{n}\hat{a} - \frac{b}{n}\sum_{i=1}^{n} x_i = 0\end{aligned}$$

$$\overline{\mathbf{y}} - \hat{a} - b\overline{\mathbf{x}} = 0 \Longleftrightarrow \hat{a} = \overline{\mathbf{y}} - b\overline{\mathbf{x}}$$

Minimisation de la somme des carrés des erreurs $f(a,b)$ par rapport au paramètre b :

$$\begin{aligned}\frac{\partial f(a,b)}{\partial b} &= \frac{\partial}{\partial b}\sum_{i=1}^{n}\left(y_i - a - bx_i\right)^2 \\ &= \sum_{i=1}^{n}\frac{\partial}{\partial b}\left[(y_i-\overline{\mathbf{y}}) - b(x_i-\overline{\mathbf{x}})\right]^2 \\ &= \sum_{i=1}^{n} -2(x_i-\overline{\mathbf{x}})\left[(y_i-\overline{\mathbf{y}}) - b(x_i-\overline{\mathbf{x}})\right]\end{aligned}$$

$$\left.\frac{\partial f(a,b)}{\partial b}\right|_{b=\hat{b}} = 0$$

$$\sum_{i=1}^{n}(x_i-\overline{\mathbf{x}})\left[(y_i-\overline{\mathbf{y}}) - \hat{b}(x_i-\overline{\mathbf{x}})\right] = 0$$

$$\sum_{i=1}^{n}(x_i-\overline{\mathbf{x}})(y_i-\overline{\mathbf{y}}) - \sum_{i=1}^{n}\hat{b}(x_i-\overline{\mathbf{x}})^2 = 0$$

$$\hat{b} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{\mathbf{x}})(y_i-\overline{\mathbf{y}})}{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{\mathbf{x}})^2} \Longleftrightarrow \hat{b} = \frac{cov(\mathbf{x},\mathbf{y})}{V(\mathbf{x})}$$

Théorème :

Les estimateurs par la méthode des moindres carrés ordinaires des paramètres a et b d'une régression linéaire simple $\mathbf{y} = a + b\mathbf{x} + \varepsilon$ sont respectivement données par :

$$\hat{a} = \overline{\mathbf{y}} - b\overline{\mathbf{x}} \qquad \hat{b} = \frac{cov(\mathbf{x}, \mathbf{y})}{V(\mathbf{x})}$$

Remarque :

Les estimateurs par la méthode des moindres carrés ordinaires des paramètres β_i , $i = 0, 1, \cdots, k$ d'une régression linéaire multiple $\mathbf{y} = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \cdots + \beta_k\mathbf{x}_k + \varepsilon$ peuvent également être obtenus de la même manière que précédemment, mais les calculs seront longs et fastidieux. Dans ce cas, on passe à une écriture matricielle de la régression linéaire multiple pour tirer avantage des simplicités des calculs offertes par l'algèbre linéaire.

Application : Consommation (C) et PIB (R) aux USA entre 1982 et 1996 en milliards de dollars

C	R
3081.5	4620.3
3240.6	4803.7
3407.6	5140.1
3566.5	5323.5
3708.7	5487.7
3822.3	5649.5
...	...
4105.8	6079.4
4219.8	6244.4
4343.6	6389.6
4486.0	6610.7
4595.3	6742.1
4714.1	6928.4

$$\overline{C} = \frac{3081.5 + 3240.6 + \cdots + 4714.1}{15} = 3964.087$$

$$\overline{R} = \frac{4620.3 + 4803.7 + \cdots + 6928.4}{15} = 5872.193$$

$$V(C) = \frac{3081.5^2 + 3240.6^2 + \cdots + 4714.1^2}{15} - (3964.087)^2 = 223783.728$$

$$V(R) = \frac{4620.3^2 + 4803.7^2 + \cdots + 6928.4^2}{15} - (5872.193)^2 = 447738.633$$

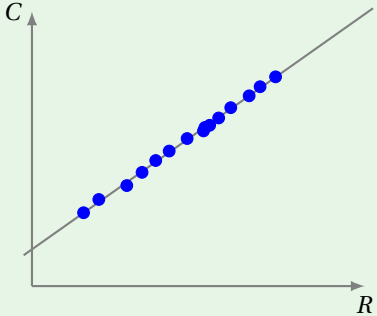
$$\overline{C.R} = \frac{3081.5 \times 4620.3 + \cdots + 4714.1 \times 6928.4}{15} = 23594169.473$$

$$cov(C,R) = 23594169.473 - 3964.087 \times 5872.193 = 316286.176$$

$$\rho(C,R) = \frac{316286,176}{\sqrt{223783.728}\sqrt{447738.633}} = 0.999$$

Application : Consommation (C) et PIB (R) aux USA entre 1982 et 1996 en milliards de dollars

Estimation des paramètres du modèle de consommation keynésien $C = C_0 + c.R + \varepsilon$:



$\hat{C} = -184.08 + 0.7064R$

- Propension marginale à consommer c :

$$\hat{c} = \frac{cov(C,R)}{V(R)} = \frac{316286.176}{447738.633} = 0.7064$$

- Consommation incompressible C_0 :

$$\hat{C}_0 = \overline{C} - \hat{c}.\overline{R} = 3964.087 - 0.7064 \times 5872.193 = -184.08$$

Chapitre 2 :

Les fondements algébriques de l'économétrie

1. Rappels de calcul matriciel
2. La méthode des moindres carrés ordinaires

Les fondements algébriques de l'économétrie

Section 1 :

Rappels de calcul matriciel

1. Matrices et vecteurs
2. Addition et soustraction de matrices
3. Multiplications matricielles
4. Transposée d'une matrice
5. Déterminant d'une matrice carrée
6. Inverse d'une matrice carrée
7. Valeurs et vecteurs propres d'une matrice carrée

Sous-section 1 :

Matrices et vecteurs

Les fondements algébriques de l'économétrie

Rappels de calcul matriciel : Matrices et vecteurs

Matrice

Une *matrice* de taille $n \times m$ est un tableau de nombres, appelés *éléments* ou *coefficients*, rangés en n lignes horizontales et m colonnes verticales :

$$A_{n\times m} \equiv \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{im} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nm} \end{bmatrix} \equiv \left[a_{ij}\right]_{n\times m}$$

Notation

Une matrice est désignée par une lettre majuscule en précisant parfois sa taille $A_{n\times m}$. Ses coefficients sont notés par la lettre minuscule correspondante "a" affectée des indices de la ligne i et de la colonne j sur lesquelles ils se situent.

Vecteur ligne

Un *vecteur ligne* $\mathbf{x}$ de taille m est une matrice composée d'une seule ligne et de m colonnes, c'est à dire : $\mathbf{x} = [x_1 \;\; x_2 \;\; \cdots \;\; x_m]$.

Vecteur colonne

Un *vecteur colonne* $\mathbf{x}$ de taille n est une matrice composée d'une seule colonne et de n lignes, c'est à dire :

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Remarque

Si la nature du vecteur n'est pas précisée, il s'agit d'un vecteur colonne.

Sous-section 2 :

Addition et soustraction de matrices

Addition de matrices

L'addition des matrices $A = [a_{ij}]_{n\times m}$ et $B = [b_{ij}]_{n\times m}$ de même taille $n \times m$ est une matrice notée $A+B$, de même taille, obtenue en additionnant les éléments correspondants des matrices A et B.

Exemple

$$\left.\begin{array}{l} A = \begin{bmatrix} 2 & 5 & 7 \\ 8 & 4 & 3 \end{bmatrix} \\ \\ B = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 5 & 3 \end{bmatrix} \end{array}\right| \Rightarrow \quad A+B = \begin{bmatrix} 2+3 & 5+2 & 7+1 \\ 8+1 & 4+5 & 3+3 \end{bmatrix} = \begin{bmatrix} 5 & 7 & 8 \\ 9 & 9 & 6 \end{bmatrix}$$

Soustraction de matrices

La soustraction des matrices $A = [a_{ij}]_{n\times m}$ et $B = [b_{ij}]_{n\times m}$ de même taille $n \times m$ est une matrice notée $A - B$, de même taille, obtenue en soustrayant les éléments de la matrice B des éléments correspondants de la matrice A.

Exemple

$$\left.\begin{array}{l} A = \begin{bmatrix} 2 & 5 & 7 \\ 8 & 4 & 3 \end{bmatrix} \\ \\ B = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 5 & 3 \end{bmatrix} \end{array}\right| \Rightarrow \quad A - B = \begin{bmatrix} 2-3 & 5-2 & 7-1 \\ 8-1 & 4-5 & 3-3 \end{bmatrix} = \begin{bmatrix} -1 & 3 & 6 \\ 7 & -1 & 0 \end{bmatrix}$$

Sous-section 3 :

Multiplications matricielles

Multiplication d'une matrice par un scalaire

Le produit d'une matrice $A = [a_{ij}]$ par un scalaire α est une matrice αA (notée également $\alpha \cdot A$ ou $\alpha \times A$) obtenue en multipliant tous les coefficients de la matrice A par le même scalaire α, c'est à dire : $\alpha A = [\alpha a_{ij}]$.

Exemple

$$A = \begin{bmatrix} 2 & 5 & 4 \\ 8 & 4 & 3 \end{bmatrix} \quad \Rightarrow \quad 2A = \begin{bmatrix} 2\times 2 & 2\times 5 & 2\times 4 \\ 2\times 8 & 2\times 4 & 2\times 3 \end{bmatrix} = \begin{bmatrix} 4 & 10 & 8 \\ 16 & 8 & 6 \end{bmatrix}$$

Produit scalaire

Le *produit scalaire* de deux vecteurs de même taille $\mathbf{a} = (a_1, a_2, \cdots, a_n)$ et $\mathbf{b} = (b_1, b_2, \cdots, b_n)$ est un scalaire, noté $\langle \mathbf{a}, \mathbf{b} \rangle$ ou $\mathbf{a}.\mathbf{b}$, donné par :

$$\langle \mathbf{a}, \mathbf{b} \rangle = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

Si les vecteurs $\mathbf{a}$ et $\mathbf{b}$ sont identiques, $\langle \mathbf{a}, \mathbf{b} \rangle$ est appelé *carré scalaire*.

Exemple

$$\mathbf{a} = (\ 2\ ,\ 5\ ,\ 3\ ) \qquad \mathbf{b} = (\ 4\ ,\ 1\ ,\ 6\ )$$

$$\langle \mathbf{a}, \mathbf{b} \rangle = 2 \times 4 + 5 \times 1 + 3 \times 6 = 31$$

Vecteurs orthogonaux

Deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont *orthogonaux* (ou *perpendiculaires*) si leur produit scalaire est nul, c'est à dire : $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. Plus généralement, un ensemble de vecteurs $\mathbf{x}_1$, $\mathbf{x}_2$,... et $\mathbf{x}_k$ sont *orthogonaux*, s'ils le sont deux à deux : $\langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0$ pour $i \neq j$.

Exemple

$$\mathbf{x} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \quad \Rightarrow \quad \langle \mathbf{x}, \mathbf{y} \rangle = 1 \times 1 + 1 \times 1 + 1 \times (-2) = 0$$

$$\mathbf{y} = \begin{bmatrix} 1 & 1 & -2 \end{bmatrix} \quad \Rightarrow \quad \langle \mathbf{x}, \mathbf{z} \rangle = 1 \times 1 + 1 \times (-1) + 1 \times 0 = 0$$

$$\mathbf{z} = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \quad \Rightarrow \quad \langle \mathbf{y}, \mathbf{z} \rangle = 1 \times 1 + 1 \times (-1) + (-2) \times 0 = 0$$

Produit matriciel

Le *produit matriciel* de deux matrices $A = [a_{ij}]_{n\times p}$ et $B = [b_{ij}]_{p\times q}$ est une matrice de taille $n \times q$, notée AB, $A\cdot B$ ou $A \times B$, dont les coefficients π_{ij} sont donnés par la produit scalaire de la ligne i de la matrice A et la colonne j de la matrice B : $\pi_{ij} = a_{i1}b_{1j} + \cdots + a_{ik}b_{kj} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$

Exemple

$$A = \begin{bmatrix} 1 & 7 \\ 3 & 5 \\ 4 & 2 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$$

$$AB = \quad ?$$

$$\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \leftarrow B$$

$$A \rightarrow \begin{bmatrix} 1 & 7 \\ 3 & 5 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 15 & 21 \\ 13 & 15 \\ 8 & 6 \end{bmatrix} \leftarrow AB$$

Remarque 1 :

Le produit matriciel de deux matrices A et B est généralement non commutatif, c'est à dire $AB \neq BA$. Dans le cas exceptionnel où $AB = BA$, on dit que les matrices A et B commutent.

Exemple

$$A = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix} \quad \Rightarrow \quad AB = BA = \begin{bmatrix} 6 & 3 \\ 0 & 3 \end{bmatrix}$$

$$A = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix} \quad C = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \quad \Rightarrow \quad AC = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix} \neq \begin{bmatrix} 6 & 3 \\ 2 & 2 \end{bmatrix} = CA$$

Remarque 2 :

Si l'une des matrices A et B est nulle alors leur produit AB est par construction nul. Le contraire n'est pas nécessairement vrai : le produit AB peut être nul sans qu'aucune des matrices A et B ne le soit.

Exemple

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{O}_{2\times 2} \quad B = \begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix} \neq \mathbf{O}_{2\times 2} \quad \Rightarrow \quad AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{O}_{2\times 2}$$

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \neq \mathbf{O}_{2\times 2} \quad B = \begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix} \neq \mathbf{O}_{2\times 2} \quad \Rightarrow \quad AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{O}_{2\times 2}$$

Remarque 3 :

L'égalité des produits matriciels AB et AC n'implique pas nécessairement l'égalité de B et C.

Exemple

$$\left.\begin{array}{l} A = \begin{bmatrix} 0 & 1 \\ 0 & 4 \end{bmatrix} \\ \\ B = \begin{bmatrix} 4 & -1 \\ 5 & 3 \end{bmatrix} \\ \\ C = \begin{bmatrix} 2 & 5 \\ 5 & 3 \end{bmatrix} \end{array}\right| \Rightarrow \quad AB = AC = \begin{bmatrix} 5 & 3 \\ 20 & 12 \end{bmatrix}$$

Sous-section 4 :

Transposée d'une matrice

Transposée d'une matrice

La *matrice transposée* (ou la *transposée*) de la matrice $A = [a_{ij}]_{n\times m}$ est une matrice de taille $m \times n$, notée $A^\top$, A' ou ${}^tA$, dont les lignes et les colonnes sont respectivement les colonnes et les lignes de la matrice A, c'est à dire : $A^\top = [a_{ji}]_{m\times n}$.

Exemple

$$A = \begin{bmatrix} 2 & 5 & 4 \\ 8 & 4 & 3 \end{bmatrix} \qquad A^\top = \begin{bmatrix} 2 & 8 \\ 5 & 4 \\ 4 & 3 \end{bmatrix}$$

Transposée d'une somme (ou différence) de matrices

La transposée de la somme (respectivement, différence) des matrices A et B est égale la somme (respectivement, différence) de leurs transposées respectives $A^\top$ et $B^\top$: $(A \pm B)^\top = A^\top \pm B^\top$.

Exemple

$$A = \begin{bmatrix} 2 & 5 & 7 \\ 8 & 4 & 3 \end{bmatrix} \Rightarrow (A \pm B)^\top = \begin{bmatrix} 2\pm 3 & 5\pm 2 & 7\pm 1 \\ 8\pm 1 & 4\pm 5 & 3\pm 3 \end{bmatrix}^\top = \begin{bmatrix} 2\pm 3 & 8\pm 1 \\ 5\pm 2 & 4\pm 5 \\ 7\pm 1 & 3\pm 3 \end{bmatrix}$$

$$B = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 5 & 3 \end{bmatrix} \Rightarrow A^\top \pm B^\top = \begin{bmatrix} 2 & 8 \\ 5 & 4 \\ 7 & 3 \end{bmatrix} \pm \begin{bmatrix} 3 & 1 \\ 2 & 5 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 2\pm 3 & 8\pm 1 \\ 5\pm 2 & 4\pm 5 \\ 7\pm 1 & 3\pm 3 \end{bmatrix}$$

Transposée d'un produit matriciel

La transposée du produit des matrices A et B est égale au produit de leurs transposées $A^\top$ et $B^\top$ dans un ordre inversé : $(AB)^\top = B^\top A^\top$.

Exemple

$$A = \begin{bmatrix} 1 & 7 & 2 \\ 3 & 5 & 4 \end{bmatrix} \quad \Rightarrow \quad (AB)^\top = \begin{bmatrix} 47 & 55 \\ 41 & 47 \end{bmatrix}^\top = \begin{bmatrix} 47 & 41 \\ 55 & 47 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 0 \\ 6 & 7 \\ 2 & 3 \end{bmatrix} \quad \Rightarrow \quad B^\top A^\top = \begin{bmatrix} 1 & 6 & 2 \\ 0 & 7 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 7 & 5 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 47 & 41 \\ 55 & 47 \end{bmatrix}$$

Sous-section 5 :

Déterminant d'une matrice carrée

Déterminant d'une matrice carrée A_n

Développement selon la ligne i de la matrice A_n :

$$det(A_n) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij})$$

Développement selon la ligne i de la matrice A_n :

$$det(A_n) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij})$$

a_{ij} est l'élément de la matrice A_n qui se trouve sur sa ligne i et sa colonne j, et A_{ij} est la sous-matrice de la matrice A_n obtenue en éliminant sa ligne i et sa colonne j.

Exemple

Développement selon la ligne 1 :

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & 4 & 1 \\ 4 & 2 & 1 \end{vmatrix} = 1 \times (-1)^{1+1} \begin{vmatrix} 4 & 1 \\ 2 & 1 \end{vmatrix} + 2 \times (-1)^{1+2} \begin{vmatrix} 3 & 1 \\ 4 & 1 \end{vmatrix} + 0 \times (-1)^{1+3} \begin{vmatrix} 3 & 4 \\ 4 & 2 \end{vmatrix}$$

$$= \begin{vmatrix} 4 & 1 \\ 2 & 1 \end{vmatrix} - 2 \times \begin{vmatrix} 3 & 1 \\ 4 & 1 \end{vmatrix} = (4-2) - 2(3-4) = 4$$

Développement selon la colonne 3 :

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & 4 & 1 \\ 4 & 2 & 1 \end{vmatrix} = 0 \times (-1)^{1+3} \begin{vmatrix} 3 & 4 \\ 4 & 2 \end{vmatrix} + 1 \times (-1)^{2+3} \begin{vmatrix} 1 & 2 \\ 4 & 2 \end{vmatrix} + 1 \times (-1)^{3+3} \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 2 \\ 4 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -(2-8) + (4-6) = 4$$

Sous-section 6 :

Inverse d'une matrice carrée

Inverse d'une matrice carrée

- Si $det(A) \neq 0$, la matrice A est inversible et sa matrice inverse A^{-1} est donnée par :

$$A^{-1} = \frac{1}{det(A)} A^a$$

avec $A^a = (A^c)^\top$ et $A^c = [c_{ij}]$ et $c_{ij} = (-1)^{i+j} det(A_{ij})$ et A_{ij} sous-matrice de la matrice A obtenue en éliminant sa ligne i et sa colonne j.

- Si $det(A) = 0$, la matrice A n'est pas inversible et sa matrice inverse A^{-1} n'est pas définie.

Exemples

$$A = \begin{bmatrix} 3 & 4 \\ 2 & 5 \end{bmatrix} \Rightarrow det(A) = 7 \Rightarrow A^c = \begin{bmatrix} 5 & -2 \\ -4 & 3 \end{bmatrix} \quad A^a = \begin{bmatrix} 5 & -4 \\ -2 & 3 \end{bmatrix} \quad A^{-1} = \begin{bmatrix} \frac{5}{7} & -\frac{4}{7} \\ -\frac{2}{7} & \frac{3}{7} \end{bmatrix}$$

$$B = \begin{bmatrix} 3 & 3 \\ 2 & 2 \end{bmatrix} \Rightarrow det(B) = 0 \rightarrow B \text{ est singulière}$$

Sous-section 7 :

Valeurs et vecteurs propres d'une matrice carrée

Valeurs et vecteurs propres d'une matrice carrée

Un vecteur non nul $\mathbf{v}$ est un *vecteur propre* (ou un *vecteur caractéristique*) d'une matrice carrée A s'il existe un scalaire λ tel que $A\mathbf{v} = \lambda\mathbf{v}$. Le scalaire λ est alors une *valeur propre* (ou une *valeur caractéristique*) de la matrice A.

Exemple

$$A = \begin{bmatrix} 1 & 4 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} \qquad \mathbf{v}_1 = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

$$A\mathbf{v}_1 = \begin{bmatrix} 8 \\ 2 \\ 0 \end{bmatrix} = 2\mathbf{v}_1 \qquad A\mathbf{v}_2 = \begin{bmatrix} 9 \\ 6 \\ 6 \end{bmatrix} = 3\mathbf{v}_2 \qquad A\mathbf{v}_3 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} = 1\mathbf{v}_3$$

Section 2 :

La méthode des moindres carrés ordinaires

1. Régression linéaire multiple
2. Écriture matricielle de la régression linéaire multiple
3. Estimateurs des moindres carrés ordinaires
4. Cas particulier : Estimateurs des paramètres d'une régression linéaire simple
5. Existence et unicité
6. Application

Sous-section 1 :

Régression linéaire multiple

Les fondements algébriques de l'économétrie

La méthode des moindres carrés ordinaires : Régression linéaire multiple

Régression linéaire multiple :

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \cdots + \beta_k \mathbf{x}_k + \varepsilon$$

Régression linéaire multiple pour chaque observation i , $i = 1, 2, \cdots, n$:

$$\begin{aligned}
y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \cdots + \beta_k x_{k1} + \varepsilon_1 \\
y_2 &= \beta_0 + \beta_1 x_{12} + \beta_2 x_{22} + \cdots + \beta_k x_{k2} + \varepsilon_2 \\
&\vdots \\
y_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i \\
&\vdots \\
y_n &= \beta_0 + \beta_1 x_{1n} + \beta_2 x_{2n} + \cdots + \beta_k x_{kn} + \varepsilon_n
\end{aligned}$$

Sous-section 2 :

Écriture matricielle de la régression linéaire multiple

Les fondements algébriques de l'économétrie

La méthode des moindres carrés ordinaires : Écriture matricielle de la régression linéaire multiple

Écriture matricielle de la régression linéaire multiple :

$$y = X\beta + \varepsilon$$

avec :

$$\underset{n\times 1}{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix} \quad \underset{n\times(k+1)}{X} = \begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & x_{1i} & x_{2i} & \cdots & x_{ki} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{pmatrix} \quad \underset{(k+1)\times 1}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \quad \underset{n\times 1}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Erreurs :

$$\varepsilon = y - X\beta$$

Sous-section 3 :

Estimateurs des moindres carrés ordinaires

Les fondements algébriques de l'économétrie

La méthode des moindres carrés ordinaires : Estimateurs des moindres carrés ordinaires

Somme des carrés des erreurs :

$$\sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon^\top \varepsilon = (y - X\beta)^\top (y - X\beta)$$

Estimateurs des moindres carrés ordinaires :

$$\hat{\beta} = \arg\min_{\beta} (y - X\beta)^\top (y - X\beta)$$

Rappels :

- Dérivée du produit scalaire des vecteurs $\mathbf{u}(\mathbf{x})$ et $\mathbf{v}(\mathbf{x})$ par rapport au vecteur $\mathbf{x}$:

$$\frac{\partial \mathbf{u}^\top \mathbf{v}}{\partial \mathbf{x}} = \frac{\partial \mathbf{u}}{\partial \mathbf{x}}.\mathbf{v} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}}.\mathbf{u}$$

- Dérivée du produit matriciel de la matrice $\mathbf{A}$ et du vecteur $\mathbf{x}$ par rapport au vecteur $\mathbf{x}$:

$$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}^\top$$

Les fondements algébriques de l'économétrie

La méthode des moindres carrés ordinaires : Estimateurs des moindres carrés ordinaires

Minimisation de la somme des carrés des erreurs $(y - X\beta)^\top (y - X\beta)$ par rapport à β :

$$\frac{\partial \varepsilon^\top \varepsilon}{\partial \beta} = \frac{\partial \varepsilon}{\partial \beta}\varepsilon + \frac{\partial \varepsilon}{\partial \beta}\varepsilon = 2\frac{\partial \varepsilon}{\partial \beta}\varepsilon$$

$$\frac{\partial \varepsilon}{\partial \beta} = \frac{\partial (y - X\beta)}{\partial \beta} = \frac{\partial y}{\partial \beta} - \frac{\partial X\beta}{\partial \beta} = \mathbf{0} - X^\top$$

$$\frac{\partial \varepsilon^\top \varepsilon}{\partial \beta} = -2X^\top (y - X\beta)$$

$$\left.\frac{\partial \varepsilon^\top \varepsilon}{\partial \beta}\right|_{\beta=\hat{\beta}} = 0 \Longleftrightarrow X^\top (y - X\hat{\beta}) = 0 \Longleftrightarrow X^\top y = X^\top X\hat{\beta}$$

$$\hat{\beta} = (X^\top X)^{-1} X^\top y$$

Théorème :

L'estimateur par la méthode des moindres carrés ordinaires des paramètres $\beta = (\beta_0, \beta_1, \cdots, \beta_k)^\top$ d'une régression linéaire multiple $\mathbf{y} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \cdots + \beta_k \mathbf{x}_k + \varepsilon$ est donné par :

$$\hat{\beta} = (X^\top X)^{-1} X^\top y$$

Remarque :

Les estimateurs par la méthode des moindres carrés ordinaires des paramètres a et b d'une régression linéaire simple $\mathbf{y} = a + b\mathbf{x} + \varepsilon$ peuvent également être obtenus de la même manière que précédemment à partir de l'écriture matricielle $y = X\beta + \varepsilon$ de la régression linéaire simple.

Sous-section 4 :

Cas particulier : Estimateurs des paramètres d'une régression linéaire simple

Cas particulier : Estimateurs des paramètres d'une régression linéaire simple

Les estimateurs par la méthode des moindres carrés ordinaires des paramètres a et b d'une régression linéaire simple $\mathbf{y} = a + b\mathbf{x} + \varepsilon$ peuvent également être obtenus de la même manière que précédemment à partir de l'écriture matricielle $y = X\beta + \varepsilon$ de la régression linéaire simple, avec :

$$\underset{n\times 1}{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix} \quad \underset{n\times 2}{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_i \\ \vdots & \ddots \\ 1 & x_n \end{pmatrix} \quad \underset{2\times 1}{\beta} = \begin{pmatrix} a \\ b \end{pmatrix} \quad \underset{n\times 1}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \hat{\beta} = (X^\top X)^{-1} X^\top y = \begin{pmatrix} \overline{y} - \hat{b}\overline{x} \\ \frac{cov(x,y)}{v(x)} \end{pmatrix}$$

Sous-section 5 :

Existence et unicité

Sous-section 6 :

Applications

Application 1 : Consommation (C) et PIB (R) aux USA entre 1940 et 1950 en milliards de dollars

Années	R	C
1940	241	226
1941	280	240
1942	319	235
1943	331	245
1944	345	255
1945	340	265
1946	332	295
1947	320	300
1948	339	305
1949	338	315
1950	371	325

Modèle de consommation keynésien **sans** prise en compte de l'impact de la deuxième guerre mondiale (une régression linéaire simple) :

$$C = \beta_0 + \beta_1.R + \varepsilon = X\beta + \varepsilon$$

Estimation par la méthode des moindres carrés ordinaires selon l'approche statistique

Années	R	C
1940	241	226
1941	280	240
1942	319	235
1943	331	245
1944	345	255
1945	340	265
1946	332	295
1947	320	300
1948	339	305
1949	338	315
1950	371	325

$$\overline{R} = \frac{241 + 280 + \cdots + 371}{11} = 323.273$$

$$\overline{C} = \frac{226 + 240 + \cdots + 325}{11} = 273.273$$

$$V(R) = \frac{241^2 + 280^2 + \cdots + 371^2}{11} - (273.273)^2 = 1118.198347$$

$$V(C) = \frac{226^2 + 240^2 + \cdots + 325^2}{11} - (323.273)^2 = 1147.107438$$

$$\overline{C.R} = \frac{226 \times 241 + \cdots + 325 \times 371}{11} = 89107.364$$

$$cov(C, R) = 89107.364 - 273.273 \times 323.273 = 765.744$$

$$\rho(C, R) = \frac{765.744}{\sqrt{1147.107438}\sqrt{1118.198347}} = 0.676$$

Estimation par la méthode des moindres carrés ordinaires selon l'approche statistique

Années	R	C
1940	241	226
1941	280	240
1942	319	235
1943	331	245
1944	345	255
1945	340	265
1946	332	295
1947	320	300
1948	339	305
1949	338	315
1950	371	325

$$\hat{\beta}_0 = \overline{C} - \hat{\beta}_1 \times \overline{R} = 273.273 - 0.6848 \times 323.273 = 51.8951$$

$$\hat{\beta}_1 = \frac{cov(C,R)}{V(R)} = \frac{765.744}{1118.198347} = 0.6848$$

Estimation par la méthode des moindres carrés ordinaires selon l'approche algébrique

$$\underset{11\times 1}{C} = \begin{pmatrix} 226 \\ 240 \\ 235 \\ 245 \\ 255 \\ 265 \\ 295 \\ 300 \\ 305 \\ 315 \\ 325 \end{pmatrix} \quad \underset{11\times 2}{X} = \begin{pmatrix} 1 & 241 \\ 1 & 280 \\ 1 & 319 \\ 1 & 331 \\ 1 & 345 \\ 1 & 340 \\ 1 & 332 \\ 1 & 320 \\ 1 & 339 \\ 1 & 338 \\ 1 & 371 \end{pmatrix} \quad \underset{2\times 1}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \quad \underset{11\times 1}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \end{pmatrix}$$

$$\hat{\beta} = (X^\top X)^{-1} X^\top C = \begin{pmatrix} 51.8951 \\ 0.6848 \end{pmatrix}$$

Application 2 : Consommation (C) et PIB (R) aux USA entre 1940 et 1950 en milliards de dollars

Années	R	C	G
1940	241	226	0
1941	280	240	0
1942	319	235	1
1943	331	245	1
1944	345	255	1
1945	340	265	1
1946	332	295	0
1947	320	300	0
1948	339	305	0
1949	338	315	0
1950	371	325	0

Modèle de consommation keynésien **avec** prise en compte de l'impact de la deuxième guerre mondiale (une régression linéaire multiple) :

$$C = \beta_0 + \beta_1.R + \beta_2.G + \varepsilon = X\beta + \varepsilon$$

Estimation par la méthode des moindres carrés ordinaires selon l'approche algébrique

$$\underset{11\times 1}{C} = \begin{pmatrix} 226 \\ 240 \\ 235 \\ 245 \\ 255 \\ 265 \\ 295 \\ 300 \\ 305 \\ 315 \\ 325 \end{pmatrix} \quad \underset{11\times 3}{X} = \begin{pmatrix} 1 & 241 & 0 \\ 1 & 280 & 0 \\ 1 & 319 & 1 \\ 1 & 331 & 1 \\ 1 & 345 & 1 \\ 1 & 340 & 1 \\ 1 & 332 & 0 \\ 1 & 320 & 0 \\ 1 & 339 & 0 \\ 1 & 338 & 0 \\ 1 & 371 & 0 \end{pmatrix} \quad \underset{3\times 1}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} \quad \underset{11\times 1}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \end{pmatrix}$$

$$\hat{\beta} = (X^\top X)^{-1} X^\top C = \begin{pmatrix} 14.4954 \\ 0.8575 \\ -50.6897 \end{pmatrix}$$

Chapitre 3 :

La géométrie des moindres carrés ordinaires

1. La géométrie du calcul matriciel
2. La géométrie des moindres carrés ordinaires

La géométrie des moindres carrés ordinaires

Davidson et MacKinnon (2021), *Foundations of Econometrics, p. 52*

> *"Une fois que l'on a une compréhension approfondie de la géométrie impliquée dans les moindres carrés ordinaires, on peut souvent s'épargner de nombreuses lignes d'algèbre fastidieuses par un simple argument géométrique."*

Section 1 :

La géométrie du calcul matriciel

1. La notion de vecteur en géométrie
2. Les coordonnées cartésiennes d'un vecteur
3. La géométrie de l'addition et de la soustraction des vecteurs
4. La géométrie des multiplications matricielles
5. La géométrie du déterminant d'une matrice
6. La géométrie des valeurs et vecteurs propres d'une matrice

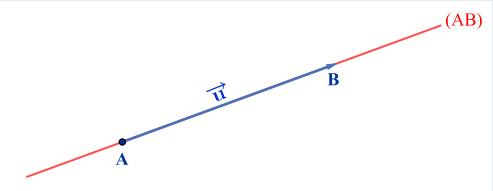
Sous-section 1 :

La notion de vecteur en géométrie

La notion de vecteur en géométrie

Un vecteur $\overrightarrow{AB}$ (ou $\overrightarrow{\mathbf{u}}$) en géométrie est un segment orienté traduisant un déplacement du point A vers le point B. Il est déterminé par :

- Une direction : la droite (AB) qui le porte ;
- Une longueur : la distance entre A et B ;
- Un sens : du point A vers le point B.



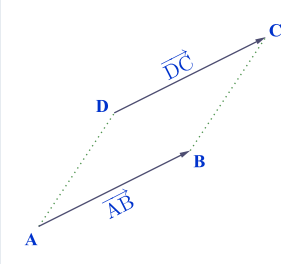
Conséquence

Si l'un des 3 éléments constitutifs d'un vecteur est modifié, le résultat est un vecteur différent.

Égalité des vecteurs

Des vecteurs $\overrightarrow{AB}$ et $\overrightarrow{DC}$ ayant la même direction, la même longueur et le même sens sont identiques.

NB : Des droites parallèles sont considérées identiques.



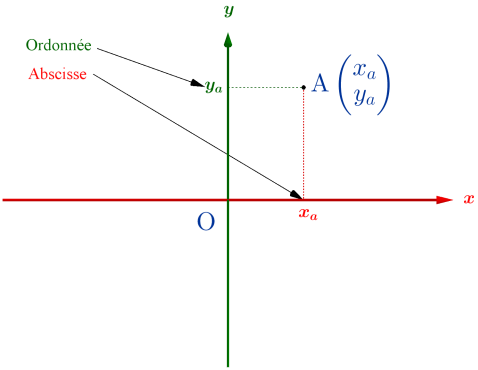
Sous-section 2 :

Les coordonnées cartésiennes d'un vecteur

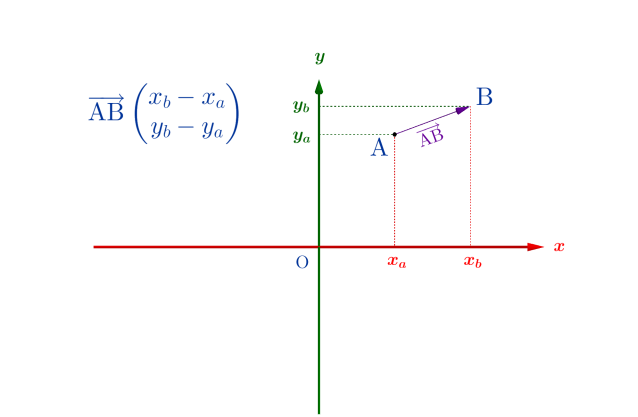
La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : Les coordonnées cartésiennes d'un vecteur

Coordonnées d'un point



Les coordonnées cartésiennes d'un vecteur

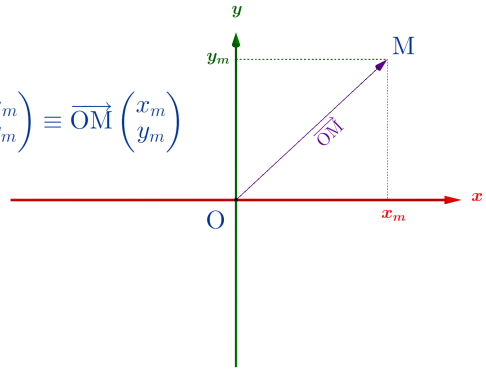


La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : Les coordonnées cartésiennes d'un vecteur

Coordonnées d'un point vs Coordonnées d'un vecteur

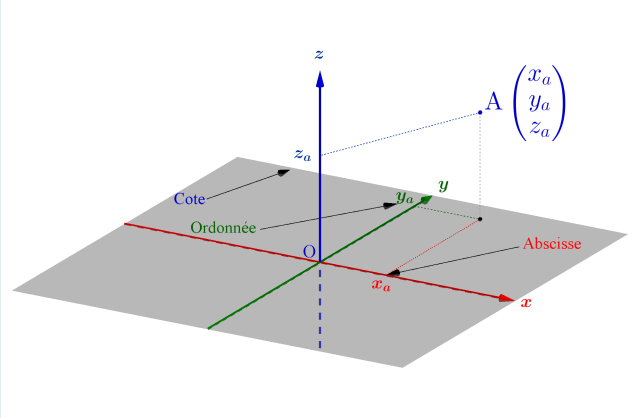
$$M\begin{pmatrix} x_m \\ y_m \end{pmatrix} \equiv \overrightarrow{\mathrm{OM}}\begin{pmatrix} x_m \\ y_m \end{pmatrix}$$



La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : Les coordonnées cartésiennes d'un vecteur

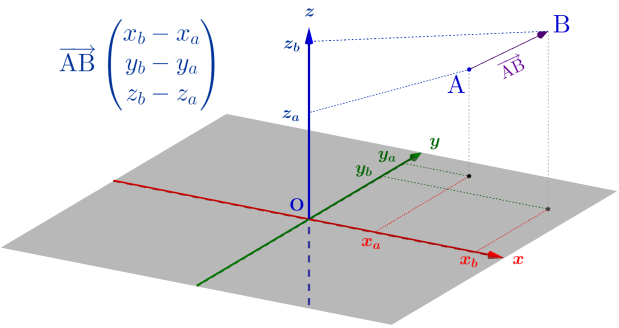
Coordonnées d'un point



La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : Les coordonnées cartésiennes d'un vecteur

Coordonnées d'un vecteur

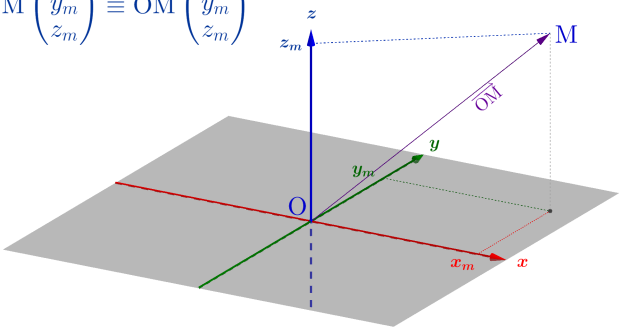


La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : Les coordonnées cartésiennes d'un vecteur

Coordonnées d'un point vs coordonnées d'un vecteur

$$\mathrm{M}\begin{pmatrix} x_m \\ y_m \\ z_m \end{pmatrix} \equiv \overrightarrow{\mathrm{OM}}\begin{pmatrix} x_m \\ y_m \\ z_m \end{pmatrix}$$



La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : Les coordonnées cartésiennes d'un vecteur

Conséquence

Les coordonnées cartésiennes $[x_1, x_2, \cdots, x_n]$ d'un point M dans un espace à n dimensions E^n sont les coordonnées cartésiennes du vecteur $\overrightarrow{OM}$ qui joint l'origine des axes O au point M. De ce fait, tout vecteur en algèbre linéaire peut être vu comme les coordonnées cartésiennes d'un point M et représenté par un vecteur $\overrightarrow{OM}$ qui relie l'origine des axes au point M.

Exemples

- Un consommateur qui consomme une quantité x du bien X et une quantité y du bien Y dispose d'un panier de consommation $P(x, y)$. Ce panier peut être représenté dans le plan (O, X, Y) par un vecteur $\overrightarrow{OP}$ reliant l'origine des axes O au point P de coordonnées (x, y).
- Un producteur utilise une quantité x du facteur travail L et une quantité y du facteur capital K pour produire une quantité $Q(x, y)$ d'un bien. Cette quantité peut être représentée dans le plan (O, L, K) par un vecteur $\overrightarrow{OQ}$ reliant l'origine des axes O au point Q de coordonnées (x, y).

Sous-section 3 :

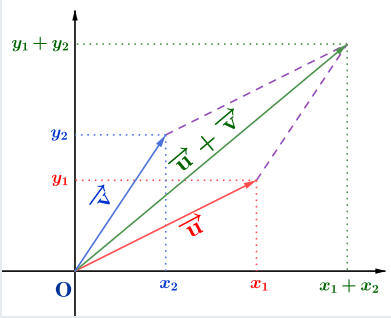
La géométrie de l'addition et de la soustraction des vecteurs

La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie de l'addition et de la soustraction des vecteurs

L'addition des vecteurs

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}$$

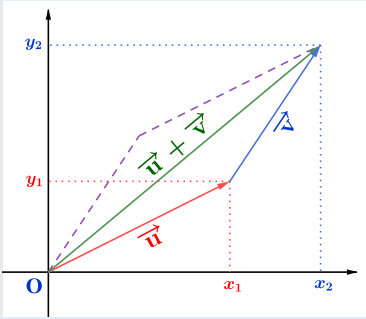


La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie de l'addition et de la soustraction des vecteurs

L'addition des vecteurs

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}$$

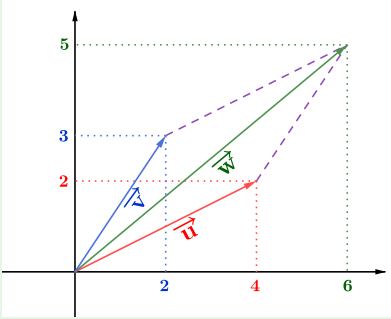


La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie de l'addition et de la soustraction des vecteurs

Exemple

$$\begin{bmatrix}4\\2\end{bmatrix}+\begin{bmatrix}2\\3\end{bmatrix}=\begin{bmatrix}4+2\\2+3\end{bmatrix}=\begin{bmatrix}6\\5\end{bmatrix}$$

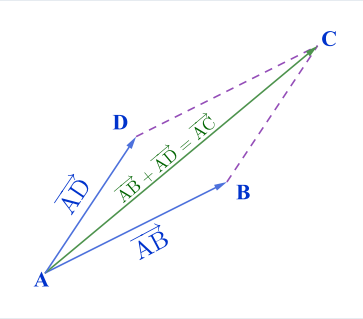


La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie de l'addition et de la soustraction des vecteurs

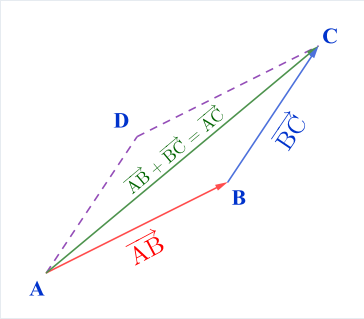
L'addition des vecteurs : Règle du parallélogramme

$$\overrightarrow{AB} + \overrightarrow{AD} = \overrightarrow{AC}$$



L'addition des vecteurs : Règle du triangle et relation de Chasles

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$$



La géométrie des moindres carrés ordinaires

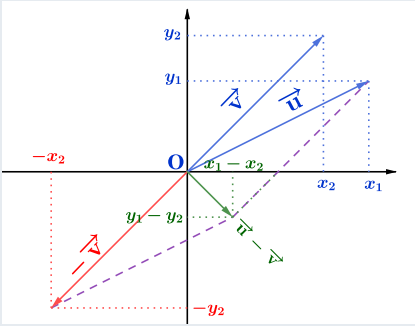
La géométrie du calcul matriciel : La géométrie de l'addition et de la soustraction des vecteurs

La soustraction des vecteurs

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} - \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ y_1 - y_2 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} -x_2 \\ -y_2 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \left(- \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right)$$

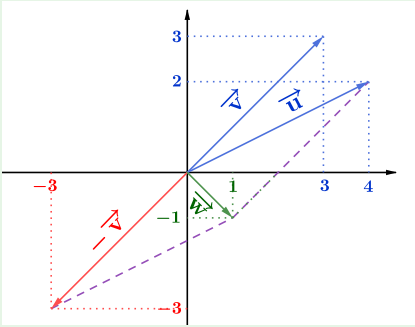


La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie de l'addition et de la soustraction des vecteurs

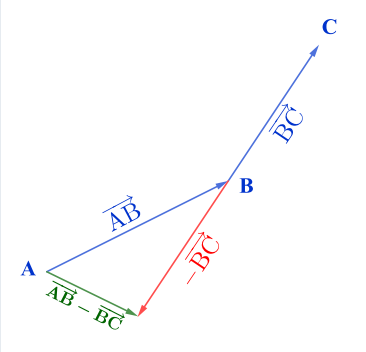
Exemple

$$\begin{aligned}\begin{bmatrix}4\\2\end{bmatrix}-\begin{bmatrix}3\\3\end{bmatrix}&=\begin{bmatrix}4-3\\2-3\end{bmatrix}\\&=\begin{bmatrix}1\\-1\end{bmatrix}\\&=\begin{bmatrix}4\\2\end{bmatrix}+\begin{bmatrix}-3\\-3\end{bmatrix}\\&=\begin{bmatrix}4\\2\end{bmatrix}+\left(-\begin{bmatrix}3\\3\end{bmatrix}\right)\end{aligned}$$



La soustraction des vecteurs : Règle du triangle

$$\begin{aligned}\overrightarrow{AB} - \overrightarrow{BC} &= \overrightarrow{AB} + (-\overrightarrow{BC}) \\ &= \overrightarrow{AB} + \overrightarrow{CB}\end{aligned}$$



Sous-section 4 :

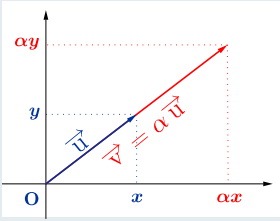
La géométrie des multiplications matricielles

La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie des multiplications matricielles

La multiplication d'un vecteur par un scalaire

$$\alpha \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha x \\ \alpha y \end{bmatrix}$$

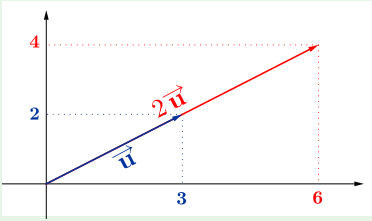


La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie des multiplications matricielles

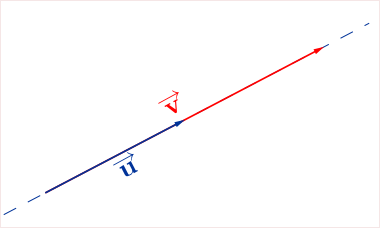
Exemple

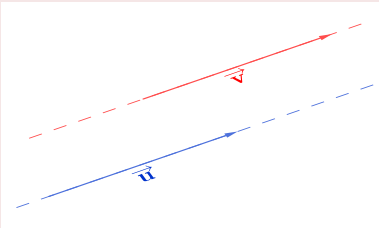
$$2 \times \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \times 3 \\ 2 \times 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$



Remarque

La multiplication d'un vecteur $\vec{\mathbf{u}}$ par un scalaire α est un vecteur $\vec{\mathbf{v}}$ qui a la même direction que le vecteur $\vec{\mathbf{u}}$. On dit que les vecteurs $\vec{\mathbf{u}}$ et $\vec{\mathbf{v}}$ sont **colinéaires**.





La géométrie des moindres carrés ordinaires

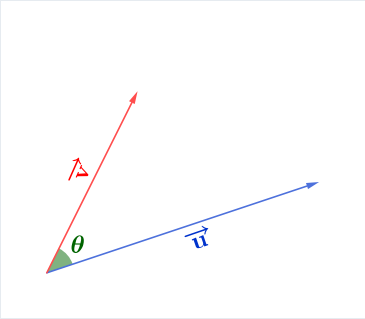
La géométrie du calcul matriciel : La géométrie des multiplications matricielles

Le produit scalaire de deux vecteurs

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\mathbf{u}^\top \mathbf{v} = \sum_{i=1}^{n} u_i . v_i$$

$$\langle \vec{\mathbf{u}}, \vec{\mathbf{v}} \rangle = \| \vec{\mathbf{u}} \| . \| \vec{\mathbf{v}} \| . \cos(\theta)$$



La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie des multiplications matricielles

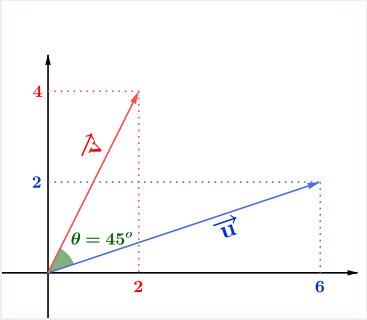
Exemple

$$\mathbf{u} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$\mathbf{u}^\top \mathbf{v} = 6 \times 2 + 2 \times 4 = 20$$

$$\langle \vec{\mathbf{u}}, \vec{\mathbf{v}} \rangle = \|\vec{\mathbf{u}}\| . \|\vec{\mathbf{v}}\| . \cos(\theta) = \sqrt{40} \times \sqrt{20} \times \frac{\sqrt{2}}{2} = 20$$



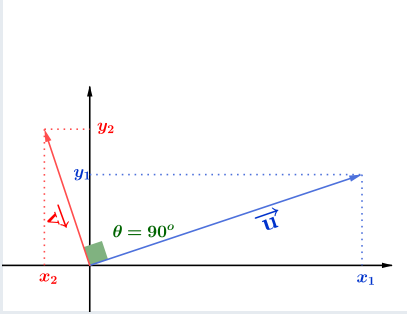
Orthogonalité et indépendance linéaire des vecteurs

$$\mathbf{u} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

$$\mathbf{u}^\top \mathbf{v} = x_1 x_2 + y_1 y_2 = 0$$

$$\langle \vec{\mathbf{u}}, \vec{\mathbf{v}} \rangle = \|\vec{\mathbf{u}}\| . \|\vec{\mathbf{v}}\| . \cos(\theta) = \|\vec{\mathbf{u}}\| . \|\vec{\mathbf{v}}\| . \cos\left(\frac{\pi}{2}\right) = 0$$



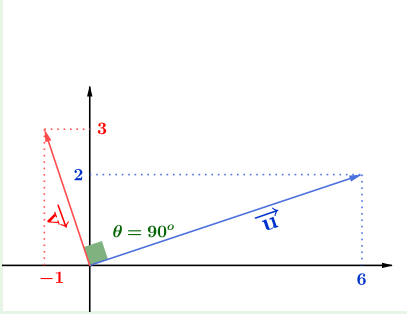
Exemple

$$\mathbf{u} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

$$\mathbf{u}^\top \mathbf{v} = 6 \times (-1) + 2 \times 3 = 0$$

$$\langle \vec{\mathbf{u}}, \vec{\mathbf{v}} \rangle = \|\vec{\mathbf{u}}\| . \|\vec{\mathbf{v}}\| . \cos(\theta) = \|\vec{\mathbf{u}}\| . \|\vec{\mathbf{v}}\| . \cos\left(\frac{\pi}{2}\right) = 0$$



La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel

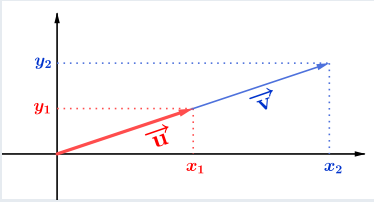
Colinéarité et dépendance linéaire des vecteurs

$$\mathbf{u} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \alpha y_1 \end{bmatrix}$$

$$\mathbf{u}^\top \mathbf{v} = x_1 x_2 + y_1 y_2 = \alpha(x_1^2 + y_1^2)$$

$$\langle \vec{\mathbf{u}}, \vec{\mathbf{v}} \rangle = \|\vec{\mathbf{u}}\| . \|\vec{\mathbf{v}}\| . \cos(0) = \alpha(x_1^2 + y_1^2)$$



La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel

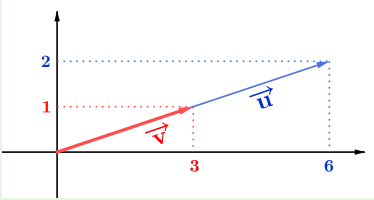
Exemple

$$\mathbf{u} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{u}^\top \mathbf{v} = 6 \times 3 + 2 \times 1 = 20$$

$$\langle \vec{\mathbf{u}}, \vec{\mathbf{v}} \rangle = \|\vec{\mathbf{u}}\| . \|\vec{\mathbf{v}}\| . \cos(0) = \sqrt{40}\sqrt{10} = 20$$



Norme d'un vecteur et théorème de Pythagore

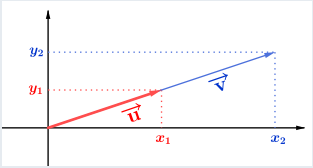
$$\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\mathbf{u}^\top \mathbf{u} = x.x + y.y = x^2 + y^2$$

$$\langle \vec{\mathbf{u}}, \vec{\mathbf{u}} \rangle = \| \vec{\mathbf{u}} \| . \| \vec{\mathbf{u}} \| . \cos(0) = x^2 + y^2$$

$$\| \vec{\mathbf{u}} \|^2 = x^2 + y^2$$

$$\| \vec{\mathbf{u}} \| = \sqrt{x^2 + y^2}$$



La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie des valeurs et vecteurs propres d'une matrice

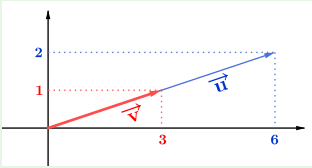
Exemple

$$\mathbf{u} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{u}^\top \mathbf{v} = 6 \times 3 + 2 \times 1 = 20$$

$$\langle \vec{\mathbf{u}}, \vec{\mathbf{v}} \rangle = \| \vec{\mathbf{u}} \| . \| \vec{\mathbf{v}} \| . \cos(0) = \sqrt{40}\sqrt{10} = 20$$



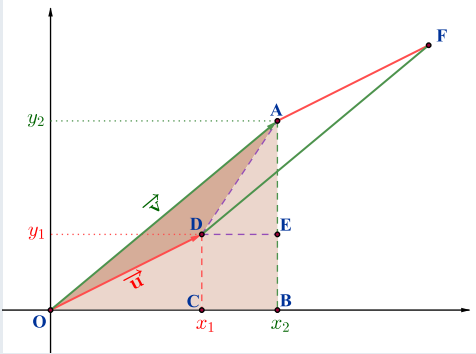
Sous-section 5 :

La géométrie du déterminant d'une matrice

Déterminant d'une matrice carrée d'ordre 2

$$A = \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} \Longrightarrow \vec{\mathbf{u}} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \ ; \ \vec{\mathbf{v}} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

Le déterminant de la matrice A est égal à l'aire $\mathscr{A}_{\text{ODFA}}$ du parallélogramme ODFA formé par les vecteurs colonnes $\vec{\mathbf{u}}$ et $\vec{\mathbf{v}}$ de la matrice A.



Exemple

$$A = \begin{bmatrix} 4 & 6 \\ 2 & 5 \end{bmatrix} \Longrightarrow \vec{\mathbf{u}} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \ ; \ \vec{\mathbf{v}} = \begin{bmatrix} 6 \\ 5 \end{bmatrix}$$

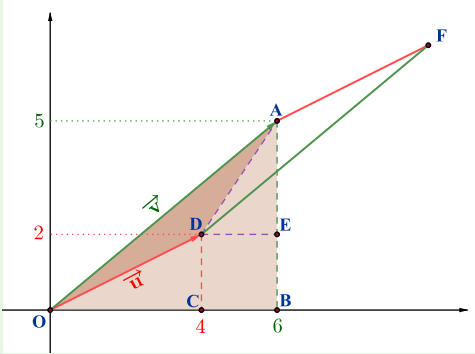
$$\det(A) = 4 \times 5 - 2 \times 6 = 8$$

$$\mathscr{A}_{\text{ODFA}} = 2\mathscr{A}_{\text{ODA}}$$

$$\mathscr{A}_{\text{ODA}} = \mathscr{A}_{\text{OBA}} - \mathscr{A}_{\text{CBED}} - \mathscr{A}_{\text{OCD}} - \mathscr{A}_{\text{AED}}$$

$$= \frac{6 \times 5}{2} - 2 \times 2 - \frac{4 \times 2}{2} - \frac{2 \times 3}{2} = 4$$

$$\mathscr{A}_{\text{ODFA}} = 2 \times 4 = 8 = \det(A)$$



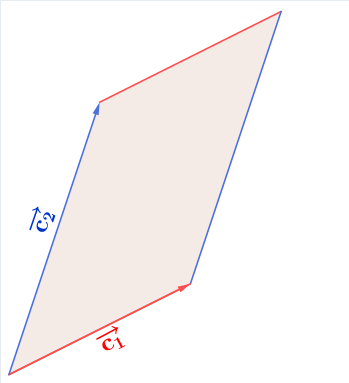
La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie du déterminant d'une matrice

Déterminant d'une matrice carrée d'ordre 2

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \Longrightarrow \overrightarrow{c_1} = \begin{bmatrix} a \\ b \end{bmatrix} \ ; \ \overrightarrow{c_2} = \begin{bmatrix} c \\ d \end{bmatrix}$$

Le déterminant de la matrice A est égal à l'aire $\mathscr{A}$ du parallélogramme formé par les vecteurs colonnes $\overrightarrow{c_1}$ et $\overrightarrow{c_2}$ de la matrice A.

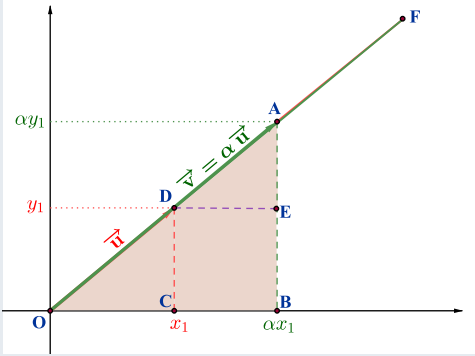


Conséquence

$$A = \begin{bmatrix} x_1 & \alpha x_1 \\ y_1 & \alpha y_1 \end{bmatrix} \Longrightarrow \vec{\mathbf{u}} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \ ; \ \vec{\mathbf{v}} = \begin{bmatrix} \alpha x_1 \\ \alpha y_1 \end{bmatrix} = \alpha \vec{\mathbf{u}}$$

$$\det(A) = x_1 \times \alpha y_1 - x_1 \times \alpha y_1 = 0$$

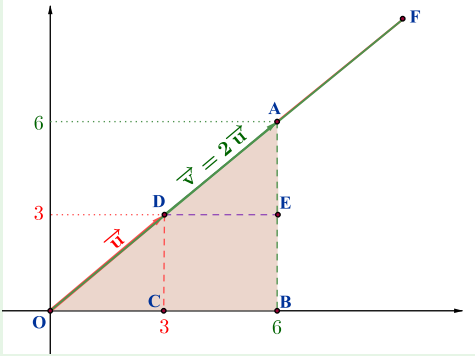
Le déterminant d'une matrice A dont les vecteurs colonnes $\vec{\mathbf{u}}$ et $\vec{\mathbf{v}}$ sont colinéaires est nul. Le parallélogramme ODFA étant aplati, son aire $\mathscr{A}_{\text{ODFA}}$ est nulle.



Exemple

$$A = \begin{bmatrix} 3 & 6 \\ 3 & 6 \end{bmatrix} \Longrightarrow \vec{\mathbf{u}} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \ ; \ \vec{\mathbf{v}} = \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \times 3 \\ 2 \times 3 \end{bmatrix} = 2\vec{\mathbf{u}}$$

$$\det(A) = 3 \times 6 - 3 \times 6 = 0$$



La géométrie des moindres carrés ordinaires

La géométrie du calcul matriciel : La géométrie du déterminant d'une matrice

Généralisation

Le déterminant d'une matrice carrée d'ordre n contenant des vecteurs lignes ou des vecteurs colonnes colinéaires est nul. Cette matrice est dite **singulière** ou **non-inversible**.

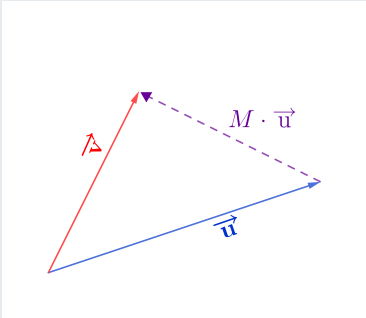
Sous-section 6 :

La géométrie des valeurs et vecteurs propres d'une matrice

Matrice de transformation

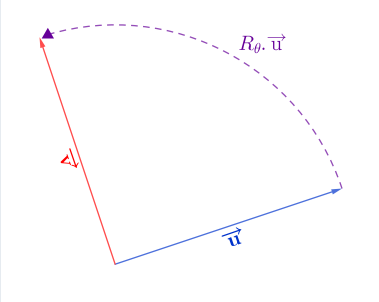
Une matrice de transformation est une matrice qui permet, à partir des coordonnées d'un point initial représentées par une matrice colonne, de trouver celles de son image par une transformation géométrique donnée. Les coordonnées de l'image sont alors obtenues en effectuant la multiplication de la matrice colonne (les coordonnées d'un point) par la matrice correspondant à la transformation géométrique concernée.

$$Mu = v$$



Matrice de rotation

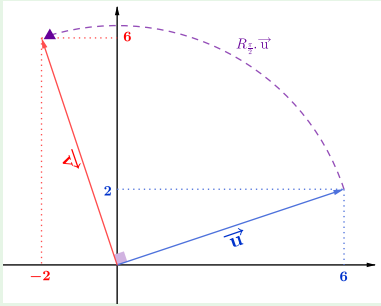
$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$



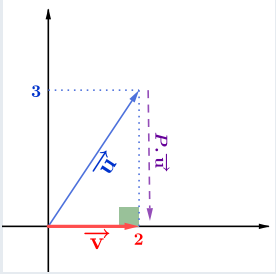
Exemple

$$R_{\frac{\pi}{2}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \; ; \; \mathbf{u} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

$$R_{\frac{\pi}{2}}\mathbf{u} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \end{bmatrix} = \mathbf{v}$$



Matrice de projection orthogonale



Valeurs et vecteurs propres d'une matrice

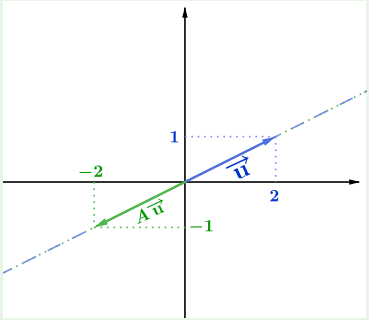
Un vecteur non nul $\mathbf{v}$ est un vecteur propre d'une matrice carrée A s'il existe un scalaire λ tel que $A\mathbf{v} = \lambda\mathbf{v}$. Le scalaire λ est alors une valeur propre de la matrice A associée au vecteur propre $\mathbf{v}$.

u est un vecteur propre de la matrice A

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \; ; \; \mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$A\mathbf{u} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix} = (-1) \times \begin{bmatrix} 2 \\ 1 \end{bmatrix} = (-1) \times \mathbf{u}$$

u est un vecteur propre de la matrice A et -1 est la valeur propre associée au vecteur propre **u**.



La géométrie des moindres carrés ordinaires

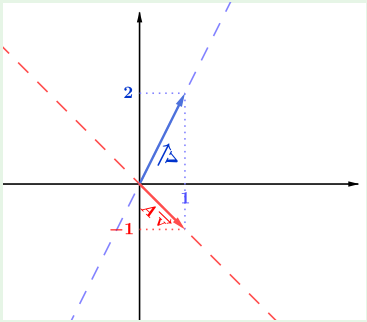
La géométrie du calcul matriciel : La géométrie des valeurs et vecteurs propres d'une matrice

$\mathbf{v}$ n'est pas un vecteur propre de la matrice A

$$A = \begin{bmatrix} 1 & 0 \\ -5 & 2 \end{bmatrix} \; ; \; \mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$A\mathbf{v} = \begin{bmatrix} 1 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \neq \lambda \mathbf{v}$$

$\mathbf{v}$ n'est pas un vecteur propre de la matrice A puisqu'il n'y a aucun scalaire non nul λ tel que $A\mathbf{v} = \lambda\mathbf{v}$.



La géométrie des moindres carrés ordinaires

Section 2 :

La géométrie des moindres carrés ordinaires

1. Écriture matricielle d'une régression linéaire
2.

Sous-section 1 :

Écriture matricielle d'une régression linéaire