

# Preuve des Moindres Carrés Ordinaires (MCO) en Format Matriciel

Simon-Pierre Boucher

Janvier 2025

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Modèle de Régression Linéaire</b>	<b>2</b>
<b>3</b>	<b>Sum of Squared Residuals</b>	<b>3</b>
<b>4</b>	<b>Minimisation de la Fonction du Sum of Squared Residuals</b>	<b>3</b>
4.1	Calcul de la Dérivée de $S$ par Rapport à $\beta$ . . . . .	3
4.2	Solution des Équations Normales . . . . .	4
<b>5</b>	<b>Conclusion</b>	<b>4</b>

# 1 Introduction

Dans cette section, nous démontrons la méthode des Moindres Carrés Ordinaires (MCO) en utilisant une approche matricielle. Cette méthode est particulièrement utile pour résoudre les problèmes de régression avec plusieurs variables explicatives (régression multiple), où les calculs matriciels permettent de simplifier le processus de minimisation.

L'objectif est de trouver les paramètres  $\beta = (\beta_1, \beta_2, \dots, \beta_k)$  du modèle de régression linéaire qui minimisent la somme des carrés des erreurs ou des résidus.

## 2 Modèle de Régression Linéaire

Le modèle de régression linéaire univarié est donné par :

$$Y_i = \beta_1 + \beta_2 X_i + \mu_i$$

où :

- $Y_i$  est la variable dépendante pour l'observation  $i$ ,
- $X_i$  est la variable indépendante (explanatoire) pour l'observation  $i$ ,
- $\beta_1$  et  $\beta_2$  sont les paramètres du modèle (intercept et pente) à estimer,
- $\mu_i$  est le terme d'erreur pour l'observation  $i$ .

Nous pouvons exprimer ce modèle de manière matricielle pour une régression multiple (c'est-à-dire avec plusieurs variables explicatives) sous la forme suivante :

$$Y = X\beta + u$$

où :

- $Y$  est un vecteur colonne  $n \times 1$  des observations de la variable dépendante,
- $X$  est une matrice  $n \times (k+1)$  des valeurs des variables explicatives, où chaque ligne contient un ensemble de variables explicatives, avec une colonne de 1 pour le terme d'intercept,
- $\beta$  est un vecteur colonne  $(k+1) \times 1$  des coefficients à estimer  $\beta_0, \beta_1, \dots, \beta_k$ ,
- $u$  est un vecteur  $n \times 1$  des erreurs ou résidus ( $u \sim N(0, \sigma^2 I)$ ).

### Hypothèses sur les erreurs :

- $E(u) = 0$ , les erreurs ont une moyenne nulle,
- $V(u) = \sigma^2 I$ , les erreurs ont une variance constante  $\sigma^2$  et sont indépendantes entre elles.



### 3 Sum of Squared Residuals

L'objectif des MCO est de minimiser la somme des carrés des résidus, qui est une mesure de l'écart total entre les valeurs observées et les valeurs prédites par le modèle. Sous forme matricielle, cette fonction de cote est donnée par :

$$S = u^\top u = (Y - X\beta)^\top (Y - X\beta)$$

Nous cherchons à minimiser  $S$ , la somme des carrés des résidus. Cela revient à minimiser l'écart entre les valeurs observées  $Y$  et les valeurs prédites  $X\beta$ .

Développons cette fonction :

$$S = (Y - X\beta)^\top (Y - X\beta) = Y^\top Y - Y^\top X\beta - \beta^\top X^\top Y + \beta^\top X^\top X\beta$$

Notez que  $Y^\top X\beta$  est un scalaire, donc il est égal à son transpose  $\beta^\top X^\top Y$ . Ainsi, nous avons :

$$S = Y^\top Y - 2\beta^\top X^\top Y + \beta^\top X^\top X\beta$$

Nous cherchons les valeurs de  $\beta$  qui minimisent cette fonction.

### 4 Minimisation de la Fonction du Sum of Squared Residuals

Pour trouver les valeurs optimales de  $\beta$ , nous devons dériver la fonction de somme des carrés des résidus  $S$  par rapport à  $\beta$  et résoudre pour  $\beta$ .

#### 4.1 Calcul de la Dérivée de $S$ par Rapport à $\beta$

Nous procédons à la dérivation de  $S$  par rapport à  $\beta$ . La dérivée de chaque terme de  $S$  est calculée comme suit :

$$\frac{\partial S}{\partial \beta} = -2X^\top Y + 2X^\top X\beta$$

Cette dérivée nous indique comment la fonction  $S$  change en fonction des paramètres  $\beta$ . Nous égalons cette dérivée à zéro pour trouver le minimum :

$$-2X^\top Y + 2X^\top X\beta = 0$$

Simplifions cette équation :

$$X^\top X\beta = X^\top Y$$

Cela nous donne les équations normales, qui sont les conditions nécessaires pour que  $S$  soit minimisé.



## 4.2 Solution des Équations Normales

Les équations normales sont :

$$X^{\top} X \beta = X^{\top} Y$$

Pour résoudre pour  $\beta$ , nous isolons  $\beta$  :

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

Cela donne l'estimateur des Moindres Carrés Ordinaires pour les paramètres  $\beta$ . Cette solution exprime les coefficients estimés comme une combinaison linéaire des variables explicatives  $X$ , pondérées par les valeurs observées  $Y$ .

## 5 Conclusion

Les Moindres Carrés Ordinaires offrent une méthode efficace pour estimer les paramètres d'un modèle de régression linéaire. En utilisant la matrice  $X$ , le vecteur  $Y$ , et les équations normales, nous pouvons obtenir les estimations des paramètres  $\beta$  qui minimisent la somme des carrés des résidus. La solution matricielle est donnée par :

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

Cet estimateur est largement utilisé en économétrie et dans d'autres domaines pour ajuster des modèles linéaires aux données.

