



60 Jours de Régression Linéaire

Jour 15 : Interactions entre variables

Sèjro Toussaint BOCO

Élève Ingénieur Data Scientist

19 décembre 2024

Plan

- 1 Définition d'interactions
- 2 Création des termes d'interaction
- 3 Tests de significativité
- 4 Interprétation des coefficients
- 5 Visualisation des interactions

Plan

- 1 Définition d'interactions
- 2 Création des termes d'interaction
- 3 Tests de significativité
- 4 Interprétation des coefficients
- 5 Visualisation des interactions

Définition des interactions

Une interaction entre variables signifie que l'effet d'une variable sur le résultat dépend du niveau d'une autre variable. En d'autres termes, l'impact d'une variable change en fonction de la présence ou du niveau d'une autre variable. Par exemple, supposons que vous étudiez l'effet de deux variables indépendantes, comme l'exercice physique (A) et le régime alimentaire (B), sur la perte de poids (variable dépendante). Une interaction entre ces variables se produirait si l'effet de l'exercice physique sur la perte de poids dépend du type de régime alimentaire suivi. il existe trois formes d'interaction, selon le type de variables indépendantes qu'elles impliquent :

- Interaction entre 2 variables quantitatives
- Interaction entre 1 variable quantitative et 1 variable catégorielle
- Interaction entre 2 variables catégorielles

Plan

- 1 Définition d'interactions
- 2 Création des termes d'interaction
- 3 Tests de significativité
- 4 Interprétation des coefficients
- 5 Visualisation des interactions

Création des termes d'interaction

Pour créer les termes d'interactions, il faut faire le produit des variables concernées entre elles. Pour deux variables X_1 et X_2 le terme d'interaction sera X_1X_2 .

Le modèle

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

passera à

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 X_2) + \epsilon$$

si X_1 et X_2 sont en interaction.

Plan

- 1 Définition d'interactions
- 2 Création des termes d'interaction
- 3 Tests de significativité**
- 4 Interprétation des coefficients
- 5 Visualisation des interactions

Tests de significativité

Le test de significativité pour un terme d'interaction vise à déterminer si l'effet conjoint de deux variables sur la réponse est statistiquement significatif.

Hypothèses :

$$\begin{cases} H_0 : \beta_3 = 0 \\ H_1 : \beta_3 \neq 0 \end{cases}$$

- Si $\beta_3 = 0$, le terme d'interaction n'a aucun effet significatif sur Y , X_1 et X_2 agissent indépendamment.
- Si $\beta_3 \neq 0$, le terme d'interaction a un effet significatif sur Y .

Plan

- 1 Définition d'interactions
- 2 Création des termes d'interaction
- 3 Tests de significativité
- 4 Interprétation des coefficients**
- 5 Visualisation des interactions

Interprétation des coefficients

Prenons l'exemple du modèle de regression suivant où X_1 et X_2 sont en interaction :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 X_2) + \epsilon$$

Dans les modèles linéaires si le coefficient de l'interaction (β_3) est significatif, cela indique que les variables X_1 et X_2 interagissent pour influencer Y . Plus précisément si :

- $(\beta_3) > 0$, alors X_1 et X_2 se renforcent mutuellement (Interaction positive)
- $(\beta_3) < 0$, alors X_1 et X_2 s'opposent ou réduisent leurs effets respectifs (Interaction négative)

Plan

- 1 Définition d'interactions
- 2 Création des termes d'interaction
- 3 Tests de significativité
- 4 Interprétation des coefficients
- 5 Visualisation des interactions

Visualisation des interactions

Même après un test statistique, il est utile de visualiser l'effet de l'interaction pour confirmer son importance pratique. On peut créer des courbes conditionnelles, des surfaces 3D, des graphiques en barres ou encore des heatmaps qui montrent comment Y varie en fonction de l'effet conjoint de X_1 et X_2 . Dans un cas pratique, cette visualisation sera exposée.