

Régression multiple: calcul des paramètres par inversion matricielle

Tableau de données :

y	x_1	x_2	----	x_m
y_1	x_{11}	x_{12}	----	x_{1m}
y_2	x_{21}	x_{23}	----	x_{2m}
----	----	----	----	----
----	----	----	----	----
----	----	----	----	----
y_n	x_{n1}	x_{n4}	----	x_{nm}

Notation matricielle de l'équation de régression : $\mathbf{y} = \mathbf{X} \mathbf{b}$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1m} \\ 1 & x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$$

L'ajout d'une colonne de "1" à la matrice \mathbf{X} permettra d'estimer l'ordonnée à l'origine.

Raisonnement (*Numerical ecology* 2012, p. 88) :

$$\mathbf{y} = \mathbf{X} \mathbf{b}$$

$$\mathbf{X}'\mathbf{y} = \mathbf{X}'\mathbf{X} \mathbf{b} \quad \Rightarrow \quad [\mathbf{X}'\mathbf{X}]^{-1}[\mathbf{X}'\mathbf{y}] = [\mathbf{X}'\mathbf{X}]^{-1} [\mathbf{X}'\mathbf{X}] \mathbf{b}$$

$$\Rightarrow \quad [\mathbf{X}'\mathbf{X}]^{-1}[\mathbf{X}'\mathbf{y}] = \mathbf{b}$$

Calcul des valeurs estimées par l'équation de régression :

$$\hat{\mathbf{y}} = \mathbf{X} \mathbf{b}$$

$$\hat{\mathbf{y}} = \mathbf{X} [\mathbf{X}'\mathbf{X}]^{-1} \mathbf{X}' \mathbf{y}$$

$$\hat{\mathbf{y}} = \mathbf{X} [\mathbf{X}'\mathbf{X}]^{-1} \mathbf{X}' \mathbf{y}$$

Projecteur

Au cours du test par permutation, on n'aura pas à recalculer le projecteur.
On calculera simplement :

$$\hat{\mathbf{y}}_{\text{permuté}} = \mathbf{X} [\mathbf{X}'\mathbf{X}]^{-1} \mathbf{X}' \mathbf{y}_{\text{permuté}}$$

$$\hat{\mathbf{y}}_{\text{permuté}} = \text{Projecteur} \mathbf{y}_{\text{permuté}}$$