

Annexe : Moindres carrés linéaires

1 Définition

Un problème de type "moindres carrés linéaires" est un problème d'optimisation d'un vecteur $\mathbf{x} \in \mathbb{R}^{D \times 1}$ où la fonction de coût est de la forme suivante :

$$L(\mathbf{x}) = \sum_{i=1}^N \|\mathbf{A}_i \mathbf{x} - \mathbf{b}_i\|_2^2, \quad (1)$$

où $\|\mathbf{v}\|_2^2 = \mathbf{v}^\top \mathbf{v}$. Remarquons que \mathbf{x} agit linéairement à l'intérieur de la norme. Les matrices \mathbf{A}_i sont de taille $M \times D$ et les vecteurs \mathbf{b}_i sont de taille $M \times 1$. Posons

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_N \end{bmatrix} \quad \text{et} \quad \mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_N \end{bmatrix}, \quad (2)$$

où \mathbf{A} est une matrice de taille $MN \times D$ et \mathbf{b} est un vecteur de taille $MN \times 1$. Ainsi :

$$L(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2. \quad (3)$$

L'objectif d'un problème de type "moindres carrés linéaires" est de minimiser $L(\mathbf{x})$ par rapport à \mathbf{x} , ce qui s'écrit :

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} L(\mathbf{x}). \quad (4)$$

2 Solution

Une manière de trouver la solution de (4) est de chercher un extremum local :

$$\frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial (\mathbf{Ax} - \mathbf{b})^\top (\mathbf{Ax} - \mathbf{b})}{\partial \mathbf{x}} = 2(\mathbf{Ax} - \mathbf{b})^\top \frac{\partial \mathbf{Ax} - \mathbf{b}}{\partial \mathbf{x}} = 2(\mathbf{Ax} - \mathbf{b})^\top \mathbf{A} = \mathbf{0}. \quad (5)$$

Nous obtenons le système linéaire de D équations ($\mathbf{A}^\top \mathbf{A}$ est de taille $D \times D$ et $\mathbf{A}^\top \mathbf{b}$ est de taille $D \times 1$) à D inconnues (\mathbf{x} est de taille $D \times 1$) suivant :

$$\mathbf{A}^\top \mathbf{Ax} = \mathbf{A}^\top \mathbf{b}. \quad (6)$$

Par conséquent, la solution d'un problème de type "moindres carrés linéaires" s'obtient en résolvant un système linéaire. Il suffit donc de construire la matrice $\mathbf{A}^\top \mathbf{A}$ et le vecteur $\mathbf{A}^\top \mathbf{b}$, et d'utiliser une méthode numérique de résolution de système linéaire (par exemple la méthode de Gauss-Seidel). En Python, il suffira d'appeler la fonction "np.linalg.solve" de la bibliothèque Numpy.