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 :

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

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

$$\mathtt{A} = \begin{bmatrix} \mathtt{A}_1 \\ \mathtt{A}_2 \\ \vdots \\ \mathtt{A}_N \end{bmatrix} \text{ et } \mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_N \end{bmatrix}, \tag{2}$$

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

$$L(\mathbf{x}) = \|\mathtt{A}\mathbf{x} - \mathbf{b}\|_2^2 \,. \tag{3}$$

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 :

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

2 Solution

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

$$\frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial \left(\mathtt{A}\mathbf{x} - \mathbf{b}\right)^\top \left(\mathtt{A}\mathbf{x} - \mathbf{b}\right)}{\partial \mathbf{x}} = 2\left(\mathtt{A}\mathbf{x} - \mathbf{b}\right)^\top \frac{\partial \mathtt{A}\mathbf{x} - \mathbf{b}}{\partial \mathbf{x}} = 2\left(\mathtt{A}\mathbf{x} - \mathbf{b}\right)^\top \mathtt{A} = \mathbf{0}. \tag{5}$$

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

$$\mathtt{A}^\top\mathtt{A}\mathbf{x} = \mathtt{A}^\top\mathbf{b}. \tag{6}$$

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 $\mathtt{A}^\top\mathtt{A}$ et le vecteur $\mathtt{A}^\top\mathbf{b}$, et d'utiliser une méthode numérique de résolution de système linéaire (par exemple la méthode de Gauss-Seidel). En MATLAB, il suffira d'appeler la fonction "mldivide".