Résolution d'un sur-système par la méthode des moindres carrés

1) Problématique :

Considérons un système ayant plus d'équations que d'inconnues et n'ayant pas de solution, comme par exemple :

$$\begin{cases} x + \ y = 1 \\ x + 2\,y = 2 \\ x - y = 3 \end{cases}$$

Un tel système est qualifié de sur-système

L'idée est alors de trouver le ou les couples (x, y) qui rendraient la quantité suivante minimale :

$$f(x,y) = (x + \ y - 1)^2 + (x + 2\,y - 2)^2 + (x - y - 3)^2$$

Cette quantité est un polynôme du second degré qui tend vers $+\infty$ quand le couple (x, y) tend vers l'infini et qui est positive. Elle possède donc un minimum absolu qu'elle atteint en un ou plusieurs points. Pour déterminer ces points, on exprime que les deux dérivées partielles y sont nulles. On résout ainsi :

$$\begin{cases} \dfrac{\partial f}{\partial x} = 2\,(x + \ y - 1) + 2\,(x + 2\,y - 2) + 2\,(x - \ y - 3) = 0 \\ \dfrac{\partial f}{\partial y} = 2\,(x + \ y - 1) + 2 \times 2\,(x + 2\,y - 2) - 2\,(x - \ y - 3) = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x + y - 1 + x + 2\,y - 2 + x - y - 3 = 0 \\ x + y - 1 + 2\,x + 4\,y - 4 - \ x + \ y + 3 = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} 3\,x + 2\,y = 6 \\ 2\,x + 6\,y = 2 \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{18 - 4} \begin{pmatrix} 6 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 6 \\ 2 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 32 \\ -6 \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} x = \dfrac{16}{7} \\ y = \dfrac{-3}{7} \end{cases}$$

La fonction f atteint donc son minimum en un point unique et en ce point nous avons :

$$\begin{cases} x + \ y = \dfrac{16}{7} - \dfrac{3}{7} = \dfrac{13}{7} \approx 1{,}9 \\ x + 2\,y = \dfrac{16}{7} - \dfrac{6}{7} = \dfrac{10}{7} \approx 1{,}4 \\ x - y = \dfrac{16}{7} + \dfrac{3}{7} = \dfrac{19}{7} \approx 2{,}7 \end{cases}$$

Ce couple ne vérifie pas le système initial mais un autre système dont le second membre n'est pas très éloigné de celui du système initial.

On dit que le que le sur-système a été résolu par la méthode des moindres carrés.

2) Formulation générale :

On appelle **sur-système**, un système de n équations à m inconnues où $m < n$. Si on note A la matrice de ce système, x le vecteur colonne formé avec les inconnues, et b le vecteur formé avec les constantes du second membre, le système s'écrit sous la forme :

$$A\,x = b$$

Notons alors $A_1, A_2, \dots, A_m$ les colonnes de la matrice A.

1er cas : La famille est libre

Autrement dit, la matrice A est de rang m.

> **Il existe alors un unique vecteur colonne $\boldsymbol{x}$ rendant la quantité suivante minimale :**
>
> $$\boldsymbol{f(x) = \|A\,x - b\|^2}$$
>
> **la norme étant la norme euclidienne sur $\mathbb{R}^{\boldsymbol{n}}$, c'est le vecteur solution du système :**
>
> $$\boldsymbol{{}^tA\,A\,x = {}^tA\,b}$$

Preuve :

Posons $x = (x_1,\, x_2, \dots, x_m)$ alors :

$$f(x) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} (a_{ij}\, x_j) - b_i \right)^2$$

Et pour $p \in [\![1\,; m]\!]$:

$$\frac{\partial f}{\partial x_p} = \sum_{i=1}^{n} 2 \left(\sum_{j=1}^{m} (a_{ij}\, x_j) - b_i \right) a_{ip} \; = 2\ {}^tA_p\,(A\,x - b)$$

Le minimum de f est atteint au point où les m dérivées partielles s'annulent, soit :

$$\forall\, p \in [\![1\,; m]\!] :\ {}^tA_p\,(A\,x - b) = 0$$

Ce qui se traduit par :

$${}^tA\,(A\,x - b) = 0$$

Soit encore :

$${}^tA\,A\,x = {}^tA\,b$$

Reste à prouver l'unicité de la solution de ce système en montrant que ${}^tA\,A$ est une matrice inversible.

Pour cela, on considère une famille orthonormale $U_1, U_2, \dots, U_m$ telle que :

$$Vect[U_1, U_2, \dots, U_m] = Vect[A_1, A_2, \dots, A_m]$$

laquelle peut être obtenue par le procédé de Schmidt.

Notons pour tout $j \in [\![1\,; m]\!]$:

$$A_j = \sum_{i=1}^{m} (p_{ij}\, U_i)$$

On a alors, en notant $P = (p_{ij})$ et U la matrice dont les colonnes sont $U_1, U_2, \dots, U_m$:

$$A = U\, P$$

Et P est inversible sinon, en raisonnant par l'absurde, on aurait l'existence d'un x non nul tel que $P\, x = 0$ donc $A\, x = U\, P\, x = 0$ et A ne serait pas de rang m.

Alors, compte tenu du fait que U est une matrice à colonnes orthogonales deux à deux et de norme 1 :

$$^{t}A\, A = \, ^{t}P \, ^{t}U\, U\, P = \, ^{t}P\, P$$

Ainsi $^{t}A\, A$ est inversible car P et ^{t}P le sont.

<u>2ème cas : La famille est liée</u>

A est donc de rang k avec $k < m$. Il existe alors une infinité de solutions rendant $f(x)$ minimale et ces solutions forment un espace affine de dimension inférieure ou égale à m.

<u>Preuve</u> :

Quitte à réordonner les colonnes de A, on peut supposer que ce sont les k premières qui sont libres, les $m-k$ suivantes s'exprimant sur les k premières par une relation du type :

$$A_j = \sum_{i=1}^{k} (p_{ij}\, A_i)$$

Soit encore, en désignant par P la matrice (p_{ij}), A' la matrice formée par les k premières colonnes de A et $A"$ la matrice formée par les $m-k$ suivantes :

$$A'' = A'P$$

Ainsi :

$$f(x) = \|A'x' + A"\, x" - b\|^2$$

$$= \|A'x' + A'P\, x" - b\|^2$$

$$= \|A'(x' + P\, x") - b\|^2$$

Posons $y = x' + P\, x"$

Alors :

$$f(x) = \|A' \, y - b\|^2$$

On est alors ramené au 1er cas, lequel assure l'existence d'un unique vecteur y_0 rendant cette quantité minimale, solution du système :

$$^{t}A' \, A' \, y = \; ^{t}A' \, b$$

Les vecteurs x rendant $f(x)$ minimale sont alors définis par :

$$x' + P \, x'' = y_0$$

Soit :

$$x' = y_0 - P \, x''$$

C'est donc un espace affine de dimension égale au rang de P, lequel est inférieur ou égal à m