Notes de cours de l'ISIMA, première année
http://www.isima.fr/~leborgne

Méthode des moindres carrés : meilleure approximation linéaire

Gilles Leborgne

31 mai 2005

Table des matières

1 Rappel de dérivation

Dans $\mathbb{R}^n$ on note $\vec{e}_i$ le i-ème vecteur de base canonique (toutes les composantes sont nulles sauf la i-ème qui vaut 1), et pour $\vec{x}=\sum_{i=1}^n x_i\vec{e}_i$ et $\vec{y}=\sum_{i=1}^n y_i\vec{e}_i$ on note $(\vec{x},\vec{y})_n=\sum_{i=1}^n x_iy_i$ le produit scalaire euclidien de $\mathbb{R}^n$ de norme associée $||\vec{x}||_n=(\vec{x},\vec{x})_n^{\frac{1}{2}}=(\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$.

Proposition 1.1 *Soit* $A=[a_{ij}]_{\substack{1\le i\le m\\1\le j\le n}}$ *une matrice* $m*n$. *Soit* $\vec{d}=\begin{pmatrix}d_1\\\vdots\\d_m\end{pmatrix}$ *un vecteur donné dans* $\mathbb{R}^m$. *Soit la fonction :*

$$f:\left\{\begin{array}{rl}\mathbb{R}^n & \to\mathbb{R},\\ \vec{x}=\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix} & \mapsto f(\vec{x})=||A.\vec{x}-\vec{d}||_m^2=(A.\vec{x}-\vec{d},A.\vec{x}-\vec{d})_m.\end{array}\right.$$

La fonction f *est dérivable, et ses dérivées partielles* $\frac{\partial f}{\partial x_i}(\vec{x})$ *sont données par :*

$$\frac{\partial f}{\partial x_i}(\vec{x})=2(A.\vec{x}-\vec{d},A.\vec{e}_i)_m. \tag{1.1}$$

Et un minimum de f *est donné en un point* $\vec{x}$ *tel que :*

$$A^tA.\vec{x}=A^t.\vec{d}, \tag{1.2}$$

système de n *équations à* n *inconnues (les* x_i*) appelées équations normales.*

En particulier si le rang de A *vaut* n *(et donc* $m\ge n$ *et* A *a plus de lignes que de colonnes) alors la matrice* A^tA *est une matrice* $n*n$ *inversible et il existe un unique vecteur* $\vec{x}$ *en lequel* f *atteint son minimum.*

Preuve. f est dérivable car quadratique en les x_i (polynôme de degré 2 en x_i). Comme :

$$\begin{aligned}||A.(\vec{x}+h\vec{e}_i)-\vec{d}||_m^2 &= (A.\vec{x}-\vec{d}+hA.\vec{e}_i, A.\vec{x}-\vec{d}+hA.\vec{e}_i)_m \\ &= ||A.\vec{x}-\vec{d}||_m^2 + 2h(A.\vec{e}_i, A.\vec{x}-\vec{d})_m + h^2||A.\vec{e}_i||_m^2,\end{aligned}$$

pour la dérivée dans direction i on a :

$$\frac{f(\vec{x}+h\vec{e}_i)-f(\vec{x})}{h} = 2(A.\vec{e}_i, A.\vec{x}-\vec{d})_m + o(1),$$

D'où (1.1).

Une condition nécessaire pour que f ait un minimum au point $\vec{x}$ est $\frac{\partial f}{\partial x_i}(\vec{x}) = 0$ pour tout $i = 1, ..., n$, soit ici $A^t(A.\vec{x}-\vec{d}) = \vec{0}$ dans $\mathbb{R}^m$, d'où (1.2).

Puis $\mathrm{Ker}A = \mathrm{Ker}(A^tA)$, voir exercice suivant 1.2. Et $\mathrm{rang}(A) = n - \dim(\mathrm{Ker}A)$ (théorème du rang). D'où si $\mathrm{rang}A = n$, on a $\mathrm{Ker}A = \{0\}$. D'où $\mathrm{Ker}(A^tA) = \{0\}$: A^tA est injective, d'où bijective (matrice carrée $n*n$). ∎

Exercice 1.2 A étant une matrice $m*n$, montrer que $\mathrm{Ker}A = \mathrm{Ker}(A^tA)$.

En déduire $\mathrm{rang}(A) = \mathrm{rang}(A^tA)$.

Réponse. Si $\vec{x} \in \mathrm{Ker}A$, i.e. si $A\vec{x} = 0$ alors $A^tA\vec{x} = 0$ d'où $\vec{x} \in \mathrm{Ker}(A^tA)$, d'où $\mathrm{Ker}A \subset \mathrm{Ker}(A^tA)$. Et si $\vec{x} \in \mathrm{Ker}(A^tA)$ alors $(A\vec{x}, A\vec{x}) = (A^tA\vec{x}, \vec{x}) = (0, \vec{x}) = 0$ et donc $A\vec{x} = 0$, donc $\vec{x} \in \mathrm{Ker}A$, donc $\mathrm{Ker}(A^tA) \subset \mathrm{Ker}A$.

Puis, A représentant une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$, on a $\mathrm{rang}(A) = n - \dim(\mathrm{Ker}A)$ (théorème du rang), et $A^t.A$ étant une matrice $n*n$ on a $\mathrm{rang}(A^tA) = n - \dim(\mathrm{Ker}A^tA)$. ∎

2 Cas 1-D

2.1 Les équations

On dispose de n points $(x_i, y_i)_{i=1,...,n}$ dans le plan $\mathbb{R}^2$ tels que $x_i \neq x_j$ pour tout $i \neq j$.

Problème initial :

$$\text{existe-t-il une droite } y = f(x) = ax+b \text{ qui passe par les points } (x_i, y_i)\,? \tag{2.1}$$

Non en général, dès que $n > 2$ (strictement plus de 2 points).

Problème modifié :

$$\text{existe-t-il une droite } y = f(x) = ax+b \text{ qui approche “au mieux” les valeurs } (x_i, y_i)\,? \tag{2.2}$$

Suivant l'interprétation du “au mieux” (= le plus faible possible en quel sens ?), il y a plusieurs réponses possibles.

Ici on considère la réponse concernant la norme des carrés :

On se donne a priori une droite $f(x) = ax + b$. On considère l'erreur $e_i = |f(x_i) - y_i|$ entre la position vertical de la mesure y_i et la position $f(x_i) = ax_i + b$ sur la droite donnée. On pose (somme des carrés) :

$$S_2(a,b) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n |f(x_i) - y_i|^2 = \sum_{i=1}^n (ax_i + b - y_i)^2, \tag{2.3}$$

somme qui dépend de la droite $y = ax + b$ choisie, i.e. qui dépend de a et de b.

Dans cette somme, les (x_i, y_i) sont donnés, et on cherche la droite, i.e. on cherche a et b, tels que $S_2(a, b)$ soit minimum (la somme des carrés des erreurs la plus faible possible). Comme S est définie sur l'espace vectoriel $\mathbb{R}^2$, la condition nécessaire de minimum est : a et b vérifient :

$$\frac{\partial S_2}{\partial a}(a,b) = 0, \qquad \frac{\partial S_2}{\partial b}(a,b) = 0. \tag{2.4}$$

Les deux équations de (2.4) donnent :

$$\begin{cases} \displaystyle\sum_{i=1}^n 2x_i(ax_i+b-y_i)=0, \\ \displaystyle\sum_{i=1}^n 2(ax_i+b-y_i)=0, \end{cases}$$

i.e. les inconnues a et b cherchées (qui donneront la droite cherchée) vérifient :

$$\begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} . \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{pmatrix}. \tag{2.5}$$

Cette méthode de calcul de a et b est appelé méthode des moindres carrés.

Proposition 2.1 *La matrice* $M = \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix}$ *est inversible, dès que* $n \geq 2$ *et au moins deux des* x_i *sont distincts. Dans ce cas (2.5) a une unique solution, i.e. la méthode des moindres carrés a une unique solution.*

Preuve. On est dans le cas $n \geq 2$ (on dispose de plus de deux points (x_i, y_i) de mesure). Le problème initial (2.1) était : trouver a et b tels que :

$$\begin{pmatrix} ax_1+b=y_1 \\ \vdots \\ ax_n+b=y_n \end{pmatrix}, \quad \text{i.e.} \quad \begin{pmatrix} x_1 & 1 \\ \vdots & \\ x_n & 1 \end{pmatrix} . \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

i.e., trouver $\vec{a} = (a, b)$ tels que :

$$A.\vec{a} = \vec{y}, \quad \text{où} \quad A = \begin{pmatrix} x_1 & 1 \\ \vdots & \\ x_n & 1 \end{pmatrix} \quad \text{et} \quad \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

(A est une matrice $n*2$ et $M = A^T.A$ est une matrice $2*2$.) Ce problème est sur-contraint quand $n \geq 3$: trop d'équations par rapport au nombre d'inconnues (ici a et b).

Mais on remarque que $\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (ax_i+b-y_i)^2 = ||A.\vec{a}-\vec{y}||^2$, i.e. que $S_2(\vec{a}) = ||A.\vec{a}-\vec{y}||^2$. Or on cherche le minimum de S_2, i.e. on cherche les $\vec{a}$ tels que, voir (1.2) :

$$A^t(A.\vec{a}-\vec{y}) = 0.$$

Et le rang de A est 2 car on a supposé au moins deux des x_i distincts (donc deux des lignes de A sont indépendantes), et donc $A^tA = M$ matrice $2*2$ est inversible, voir proposition 1.1. ∎

Remarque 2.2 Plutôt que S_2 on aurait pu introduire $S_1(a,b) = \sum_{i=1}^n |\tilde{y}_i - y_i| = \sum_{i=1}^n |ax_i+b-y_i|$. Mais cette fonction n'est pas dérivable en tout a ou en tout b, la fonction valeur absolue n'étant pas dérivable en 0. D'où le choix simple de S_2 (somme des carrés). ∎

2.2 La résolution

2.2.1 Première étape

Dans le cas 1-D, proposition 2.1, la matrice M peut être mal conditionnée : cas où les colonnes de $M = \begin{pmatrix} \sum_i x_i^2 & \sum_i x_i \\ \sum_i x_i & n \end{pmatrix}$ sont presques proportionnelles par exemple.

Si tel est le cas, une technique simple est de changer l'origine des x, et de même en y (pour la sensibilité au membre de droite), i.e. de poser :

$$x_m = \frac{\sum_i x_i}{n} \quad (\text{moyenne en } x), \quad y_m = \frac{\sum_i y_i}{n} \quad (\text{moyenne en } y),$$

puis de faire le changement d'origine :

$$\bar{x} = x - x_m, \quad \bar{y} = y - y_m.$$

Les points de mesure sont maintenant les $(\bar{x}_i = x_i - x_m, \bar{y}_i = y_i - y_m)$ et le problème à résoudre est : trouver $(\bar a, \bar b)$ t.q. :

$$\begin{pmatrix} \sum_i \bar x_i^2 & 0 \\ 0 & n \end{pmatrix} . \begin{pmatrix} \bar a \\ \bar b \end{pmatrix} = \begin{pmatrix} \sum_i \bar x_i \bar y_i \\ 0 \end{pmatrix}.$$

On obtient immédiatement $\bar b = 0$ et $\bar a = \frac{\sum_i \bar x_i \bar y_i}{\sum_i \bar x_i^2}$. D'où la droite d'équation $\bar y = \bar a \bar x$, d'où la droite en les coordonnées initiales $y = y_m + \bar a(x - x_m)$.

2.2.2 Seconde étape

L'étape ci-dessus est souvent suffisante. Sinon, pour éviter que $\sum_i \bar x_i^2$ soit trop grand, on se sert des écarts types :

$$\sigma_x = \Big(\sum_i \frac{(x_i - x_m)^2}{n}\Big)^{\frac12} = \Big(\sum_i \frac{\bar x_i^2}{n}\Big)^{\frac12}, \quad \sigma_y = \Big(\sum_i \frac{(y_i - y_m)^2}{n}\Big)^{\frac12} = \Big(\sum_i \frac{\bar y_i^2}{n}\Big)^{\frac12},$$

et on fait le changement de variables avec les variables centrées réduites :

$$\hat x = \frac{\bar x}{\sigma_x} = \frac{x - x_m}{\sigma_x}, \qquad \hat y = \frac{\bar y}{\sigma_y} = \frac{y - y_m}{\sigma_y},$$

ce qui donne $\sum_i \hat x_i^2 = n = \sum_i \hat y_i^2$. Intérêt supplémentaire : les variables $\hat x$ et $\hat y$ sont sans dimension et ne dépendent donc pas de l'unité de mesure choisie.

Donc les points à considérer sont les points $(\hat x_i = \frac{x_i - x_m}{\sigma_x}, \hat y_i = \frac{y_i - y_m}{\sigma_y})$, et le problème à résoudre est : trouver $(\hat a, \hat b)$ t.q. :

$$\begin{pmatrix} n & 0 \\ 0 & n \end{pmatrix} . \begin{pmatrix} \hat a \\ \hat b \end{pmatrix} = \begin{pmatrix} \sum_i \hat x_i \hat y_i \\ 0 \end{pmatrix}, \quad \text{d'où} \quad \begin{pmatrix} \hat a = \dfrac{\sum_i \hat x_i \hat y_i}{n} \\ \hat b = 0 \end{pmatrix}.$$

D'où la droite $\hat y = \hat a \hat x + \hat b$ de meilleure approximation. D'où, revenant aux coordonnées initiales, la droite de meilleure approximation donnée par :

$$\frac{y - y_m}{\sigma_y} = \hat a \frac{x - x_m}{\sigma_x}.$$

3 Cas 2-D

3.1 Les équations

On cherche un plan qui approche au mieux un nuage de points $((x_i, y_i), z_i)$ pour $i = 1, ..., n$. On cherche donc une fonction affine $f(x, y) = ax + by + c$ telle que l'erreur $e_i = |f(x_i, y_i) - z_i|$ soit "globalement" la plus faible possible. Comme précédemment, on va minimiser $\sum_i e_i^2$.

On se donne a priori "un plan non vertical" $z = f(x, y) = ax + by + c$ et on pose :

$$S_2(a, b, c) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (ax_i + by_i + c - z_i)^2, \tag{3.1}$$

somme qui dépend du plan $f(x, y) = ax + by + c$ donné, i.e qui dépend de a, b et c.

Et le meilleur plan sera donné par un triplet (a, b, c) tel que $S_2(a, b, c) = \min_{\tilde a, \tilde b, \tilde c} S_2(\tilde a, \tilde b, \tilde c)$ (i.e. qui minimise la somme des carrés des erreurs). Donc, si (a, b, c) existe, on a t.q. :

$$\frac{\partial S_2}{\partial a}(a, b, c) = 0, \qquad \frac{\partial S_2}{\partial b}(a, b, c) = 0, \qquad \frac{\partial S_2}{\partial c}(a, b, c) = 0. \tag{3.2}$$

Les trois équations de (3.2) donnent :

$$\begin{cases} \displaystyle\sum_{i=1}^n 2x_i(ax_i + by_i + c - z_i) = 0, \\ \displaystyle\sum_{i=1}^n 2y_i(ax_i + by_i + c - z_i) = 0, \\ \displaystyle\sum_{i=1}^n 2(ax_i + by_i + c - z_i) = 0. \end{cases}$$

On doit donc résoudre :

$$\begin{pmatrix} \sum_i x_i^2 & \sum_i x_i y_i & \sum_i x_i \\ \sum_i x_i y_i & \sum_i y_i^2 & \sum_i y_i \\ \sum_i x_i & \sum_i y_i & n \end{pmatrix} . \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_i x_i z_i \\ \sum_i y_i z_i \\ \sum_i z_i \end{pmatrix}. \tag{3.3}$$

Proposition 3.1 *La matrice* $M = \begin{pmatrix} \sum_i x_i^2 & \sum_i x_i y_i & \sum_i x_i \\ \sum_i x_i y_i & \sum_i y_i^2 & \sum_i y_i \\ \sum_i x_i & \sum_i y_i & n \end{pmatrix}$ *est inversible, dès que* $n \geq 3$ *et au moins trois des* (x_i, y_i) *sont non alignés. Et (2.5) a une unique solution, i.e. la méthode des moindres carrés a une unique solution.*

Preuve. Similaire à la démonstration précédente avec ici la matrice $A = \begin{pmatrix} x_1 & y_1 & 1 \\ \vdots & & \\ x_n & y_n & 1 \end{pmatrix}$ avec encore $M = A^T.A$.

Puis pour que A soit de rang 3 il faut $n \geq 3$ et trois lignes soient indépendantes. Quitte à renuméroter les points on regarde le déterminant :

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = \begin{vmatrix} x_1{-}x_3 & y_1{-}y_3 & 0 \\ x_2{-}x_3 & y_2{-}y_3 & 0 \\ x_3 & y_3 & 1 \end{vmatrix} = \begin{vmatrix} x_1{-}x_3 & y_1{-}y_3 \\ x_2{-}x_3 & y_2{-}y_3 \end{vmatrix} = \begin{vmatrix} (& \vec{x}_1{-}\vec{x}_3 &) \\ (& \vec{x}_2{-}\vec{x}_3 &) \end{vmatrix},$$

où $\vec{x}_i = (x_i, y_i)$. Ce déterminant est non nul dès que les vecteurs $\vec{x}_1{-}\vec{x}_3$ et $\vec{x}_2{-}\vec{x}_3$ sont indépendants, i.e. dès que les points sont non alignés dans le plan (x, y). ▟

3.2 La résolution

3.2.1 Première étape

Dans le cas 2-D, proposition 3.1, même démarche que dans le cas 1-D : on pose (valeurs moyennes) :

$$x_m = \frac{\sum_i x_i}{n}, \quad y_m = \frac{\sum_i y_i}{n}, \quad z_m = \frac{\sum_i z_i}{n},$$

Le changement de coordonnées (ici d'origine) choisit est :

$$\bar{x} = x{-}x_m, \quad \bar{y} = y{-}y_m, \quad \bar{z} = z{-}z_m.$$

On obtient donc les points $(\bar{x}_i = x_i{-}x_m, \bar{y}_i = y_i{-}y_m, \bar{z}_i = z_i{-}z_m)$ à approcher par un plan, et le problème à résoudre est : trouver $(\bar{a}, \bar{b}, \bar{c})$ t.q. :

$$\begin{pmatrix} \sum_i \bar{x}_i^2 & \sum_i \bar{x}_i \bar{y}_i & 0 \\ \sum_i \bar{x}_i \bar{y}_i & \sum_i \bar{y}_i^2 & 0 \\ 0 & 0 & n \end{pmatrix} . \begin{pmatrix} \bar{a} \\ \bar{b} \\ \bar{c} \end{pmatrix} = \begin{pmatrix} \sum_i \bar{x}_i \bar{z}_i \\ \sum_i \bar{y}_i \bar{z}_i \\ 0 \end{pmatrix}.$$

D'où immédiatement $\bar{c} = 0$, et il reste à résoudre :

$$\begin{pmatrix} \sum_i \bar{x}_i^2 & \sum_i \bar{x}_i \bar{y}_i \\ \sum_i \bar{x}_i \bar{y}_i & \sum_i \bar{y}_i^2 \end{pmatrix} . \begin{pmatrix} \bar{a} \\ \bar{b} \end{pmatrix} = \begin{pmatrix} \sum_i \bar{x}_i \bar{z}_i \\ \sum_i \bar{y}_i \bar{z}_i \end{pmatrix}.$$

La matrice $\bar{C} = \begin{pmatrix} \sum_i \bar{x}_i^2 & \sum_i \bar{x}_i \bar{y}_i \\ \sum_i \bar{x}_i \bar{y}_i & \sum_i \bar{y}_i^2 \end{pmatrix}$ est inversible d'après la proposition 3.1 (au moins trois points non alignés), en particulier son déterminant vaut $\det \bar{C} = (\sum_i \bar{x}_i^2)(\sum_i \bar{y}_i^2) - (\sum_i \bar{x}_i \bar{y}_i)^2 \neq 0$ (c'est également une conséquence de l'inégalité de Cauchy–Schwarz $(\vec{x}, \vec{y})_n \leq ||\vec{x}||_n ||\vec{y}||_n$ avec égalité uniquement dans le cas $\vec{x} // \vec{y}$ i.e. la matrice $n{*}2$ dont les colonnes sont $\vec{x}$ et $\vec{y}$ est de rang 1 et donc égalité ssi toutes les lignes sont proportionnelles à l'une d'entre elles).

D'où :

$$\begin{pmatrix} \bar{a} \\ \bar{b} \end{pmatrix} = \frac{1}{\det \bar{C}} \begin{pmatrix} \sum_i \bar{x}_i^2 & -\sum_i \bar{x}_i \bar{y}_i \\ -\sum_i \bar{x}_i \bar{y}_i & \sum_i \bar{y}_i^2 \end{pmatrix} \begin{pmatrix} \sum_i \bar{x}_i \bar{z}_i \\ \sum_i \bar{y}_i \bar{z}_i \end{pmatrix}.$$

3.2.2 Seconde étape

Si l'étape précédente n'est pas suffisante, on adimensionalise à l'aides des écarts types :

$$\sigma_x = (\sum_i \frac{(x_i - x_m)^2}{n})^{\frac{1}{2}}, \quad \sigma_y = (\sum_i \frac{(y_i - y_m)^2}{n})^{\frac{1}{2}}, \quad \sigma_z = (\sum_i \frac{(z_i - z_m)^2}{n})^{\frac{1}{2}}.$$

Le changement de coordonnées (d'origine et d'échelles) est alors :

$$\hat{x} = \frac{x - x_m}{\sigma_x}, \quad \hat{y} = \frac{y - y_m}{\sigma_y}, \quad \hat{z} = \frac{z - z_m}{\sigma_z},$$

et les points considérés sont les $(\hat{x}_i, \hat{y}_i, \hat{z}_i)$. Et le nouveau problème à résoudre est : trouver $(\hat{a}, \hat{b})$ t.q. :

$$\begin{pmatrix} n & \sum_i \hat{x}_i \hat{y}_i \\ \sum_i \hat{x}_i \hat{y}_i & n \end{pmatrix} . \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} \sum_i \hat{x}_i \hat{z}_i \\ \sum_i \hat{y}_i \hat{z}_i \end{pmatrix}.$$

La matrice $\hat{C} = \begin{pmatrix} n & \sum_i \hat{x}_i \hat{y}_i \\ \sum_i \hat{x}_i \hat{y}_i & n \end{pmatrix}$ est inversible, de déterminant $\det \hat{C} = n^2 - (\sum_i \hat{x}_i \hat{y}_i)^2 \neq 0$ (proposition 3.1 ou inégalité de Cauchy-Schwarz), et on obtient :

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \frac{1}{\det \hat{C}} \begin{pmatrix} n & -\sum_i \hat{x}_i \hat{y}_i \\ -\sum_i \hat{x}_i \hat{y}_i & n \end{pmatrix} \begin{pmatrix} \sum_i \hat{x}_i \hat{z}_i \\ \sum_i \hat{y}_i \hat{z}_i \end{pmatrix}.$$

4 Résolution par la méthode QR

Résoudre les équations normales (1.2) peut être difficile à cause du mauvais conditionnement éventuel de la matrice $A^t A$: si $\rho(A)$ est le conditionnement de la matrice A, alors $\rho(A)^2$ est le conditionnement de la matrice $A^t A$. Et si ce conditionnement est mauvais (par exemple de l'ordre de 10^7), il faut éviter de calculer $A^t A$.

La matrice A est de taille $m * n$ (et la matrice $A^t A$ de taille $n * n$). On décompose la matrice A sous la forme $A = Q.R$, avec Q matrice $m*m$ orthonormale et R matrice $m*n$ matrice triangulaire supérieure :

$$A = QR \qquad \text{où} \qquad Q^t Q = I_m \quad \text{et} \quad R = \begin{pmatrix} \mathtt{x} & \mathtt{x} & \dots & \mathtt{x} \\ 0 & \mathtt{x} & \dots & \mathtt{x} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \mathtt{x} \\ 0 & & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & \dots & 0 \end{pmatrix} = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}, \quad R^t R = R_1^t R_1,$$

où les $\mathtt{x}$ représentent des emplacements non (éventuellement) nuls, et $R_1 \in \mathbb{R}^{n^2}$ est une matrice $n * n$ triangulaire supérieure, et le dernier 0 de la formule tient lieu de la matrice nulle de taille $(m - n) * n$.

Dans le cas $m \geq n$ on pose alors $Q = (Q_1 \quad Q_2)$ avec Q_1 matrice $m*n$ et Q_2 matrice $m*(n-m)$. Et on a :

$$A = QR = (Q_1 \quad Q_2) \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = Q_1 R_1.$$

En particulier, dans la décomposition QR on se contente de calculer Q_1 et R_1. Et comme les colonnes de Q et donc de Q_1 sont des vecteurs orthonormaux on a :

$$Q_1^t Q_1 = I_n \qquad (\text{identité de } \mathbb{R}^n).$$

Proposition 4.1 *Pour $m \geq n$, A de rang n, et $\vec{d} \in \mathbb{R}^m$, le problème : trouver $\vec{v} \in \mathbb{R}^n$ tel que :*

$$A^t A \vec{v} = A^t \vec{d}, \tag{4.1}$$

a une unique solution qui vérifie :

$$R_1 \vec{v} = Q_1^t \vec{d}. \tag{4.2}$$

Preuve. Comme A est de rang n on a R_1 de rang n, et (4.2) a une unique solution. Comme $m \geq n$ et A est de rang n, A^tA est de rang n, et (4.1) a une unique solution.

Et la solution de (4.2) vérifie $R_1^t R_1 \vec{v} = R_1^t Q_1^t \vec{d}$, avec $R_1^t R_1 = R_1^t Q_1^t Q_1 R_1 = A^t A$ et $R_1^t Q_1^t = A^t$, donc est également la solution de (4.1). ∎

La matrice Q étant orthonormale, le conditionnement de la matrice R_1 est celui de la matrice A, alors que le conditionnement de la matrice A^tA est le carré du conditionnement de A. Résoudre le problème (4.2) est donc plus facile que résoudre le problème $A^tA\vec{v} = A^t\vec{d}$.

Enfin, pour programmer la décomposition de A en QR, on pourra utiliser soit la méthode de Givens (les rotations), soit la méthode de Householder (les symétries).

Références

[1] Golub G., Van Loan C. : Matrix computations. Johns Hopkins, Third edition, 1996.

[2] Spiegel M. : Statistique, cours et problèmes. MC Graw Hill, série Schaum, 2^e édition, 1993.

[3] Strang G. : Linear Algebra and its Applications. Harcourt Brace & Cie, 3rd edition, 1986.