

UNIVERSITÉ DE MOHAMED BOUDIAF- M'SILA

MEMOIRE

Présentée à la Faculté des Mathématiques et de l'Informatique

Département de Mathématiques

Pour l'obtention du diplôme de master

Spécialité : Mathématiques

Option : Mathématiques Fondamentales et Appliquées

Par :

Kebaili Nour El Imène

Intitulée :

La méthode des moindres carrés pour les équations de la physique

Soutenue publiquement le : 09/06/2015, devant le jury :

Rhmoun AZDINE MCA. Université de BBA Président

Mostefa NADIR Prof. Université de M'sila Rapporteur

Djaïdja NOUI MAA. Université de M'sila Co- Rapporteur

Lakehali BELKACEM MAA. Université de M'sila Examineur

Promotion : 2014 /2015

Remerciements

Je remercie ALLAH à tout-puissant de m'avoir le courage, la volonté et la patience de mener à terme ce présent travail.

On remercie chaleureusement à Monsieur : M .NADIR d'avoir accepté de nos encadrer, On les remercie pour ses conseils et sa disponibilité.

Merci également à Mr : DJAIDJA Noui

Un grand merci aux membres du jury pour l'honneur qu'ils nous ont fait en acceptant d'évaluer notre modeste travail.

Merci à ma famille et mes amis pour leurs soutiens et leurs encouragements, et à tous ceux qui ont, de près ou de loin, attribué et aidé à l'aboutissement de ce travail.

Table des matières

Introduction	1
1 Notions et généralités	3
1.1 Généralités sur les matrices	3
1.1.1 Les normes vectorielles et les normes matricielles	5
1.1.2 Conditionnement d'un système linéaire	6
1.2 Dérivation d'une forme quadratique	11
2 Problème de moindres carrés linéaires	13
2.1 Propriétés mathématiques du problèmes de moindres carrés	13
2.1.1 Formulation matricielle	13
2.1.2 Les équations normales	14
2.2 Méthodes de résolution du problème de moindres carrés	16
2.2.1 Résolution du problème de moindres carrés par la méthode (QR) . .	16
2.2.2 Résolution du problème de moindres carrés par décomposition de Cholesky	19
2.2.3 Résolution du problème de moindres carrés par décomposition en valeurs singulières	21
3 Résultats numériques	27
3.1 Exemple	27
Conclusion générale	31

Bibliographie	32
----------------------	-----------

Introduction

On appelle système linéaire une expression de la forme

$$Ax = b. \tag{1}$$

où $A = a_{ij}$ ($1 \leq i \leq m$, $1 \leq j \leq n$), désigne la matrice de taille $m \times n$,

$b = (b_i)$, $1 \leq i \leq m$, un vecteur colonne, $x = x_j$, le vecteur des inconnus du système.

Le système linéaire $Ax = b$ prendre trois types :

Premier type : $m = n$, i.e. le nombre d'équations égale au nombre d'inconnues, si

$\det(A) \neq 0$, le système $Ax = b$ admet une solution unique.

Deuxième type : $m < n$, i.e. le nombre d'équations inférieur au nombre d'inconnues, il existe une infinité des solutions et le système appelée système indéterminé.

Troisième type : $m > n$, i.e. le nombre d'équations supérieur au nombre d'inconnues, il n'existe pas de solutions et le système appelé système surdéterminé.

Dans ce mémoire, notre attention sera portée sur le troisième type, il est souvent rare de trouver un solution exacte de cet système, alors dans ce cas on s'intéresse à trouver une solution approchée au sens de moindres carrés d'un problème de minimisation associé à l'équation (1) donnée par :

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2. \tag{2}$$

où $\|\cdot\|_2^2$ désigne la norme matricielle euclidienne. Le problème est en général appelé un problème de moindres carrés.

Dans ce mémoire, on s'intéresse à la résolution numérique des équations physiques par la méthode des moindres carrés. Pour cette raison nous avons organisé notre travail comme suit :

Dans le premier chapitre, nous avons présenté des notions sur les matrices, conditionnement et la stabilité d'un système linéaire et ainsi le gradient d'une forme quadratique.

Le second chapitre, concerne quelques propriétés mathématiques des problèmes de moindres carrés et ainsi exposé quelques méthodes de résolution du problème de moindres carrés *QR*, *Cholesky* et développement en valeurs singulières.

Dans le troisième chapitre, on donne des résultats numériques issus de cet algorithme et qui implémentés numériquement sous MATLAB. On finira ce mémoire par une conclusion.

Chapitre 1

Notions et généralités

1.1 Généralités sur les matrices

Définition 1.1.1 Une matrice carrée A est inversible (régulière) lorsque il existe une matrice B telle que

$$AB = BA = I. \text{ (} I \text{ la matrice identité)}$$

Théorème 1.1.1 Soit A une matrice carrée, les trois propriétés sont équivalentes :

1. A est inversible.
2. $\det(A) = |A| \neq 0$.
3. Les vecteurs colonnes de A sont linéairement indépendants.

Définition 1.1.2 (Matrice symétrique) La matrice transposée d'une matrice carrée

$A \in M_n(\mathbb{R})$, telle que $A = [a_{ij}]_{1 \leq i, j \leq n}$ est $A^T = [a_{ji}]_{1 \leq i, j \leq n}$. On dit que A est symétrique si $A = A^T$.

Définition 1.1.3 Soit A une matrice carrée d'ordre n , A est dite orthogonale si

$$A^T A = A A^T = I \Leftrightarrow A^T = A^{-1}.$$

(I la matrice identité).

Définition 1.1.4 Une matrice $A \in M_n(\mathbb{R})$ est définie positive sur $\mathbb{R}^n$, si

$$\langle Ax, x \rangle > 0, \forall x \in \mathbb{R}^n, x \neq 0.$$

Définition 1.1.5 Une matrice $A \in M_n(\mathbb{R})$ est semi-définie positive sur $\mathbb{R}^n$, si

$$\langle Ax, x \rangle \geq 0, \forall x \in \mathbb{R}^n, x \neq 0.$$

Définition 1.1.6 Le rang de $A \in M_{m,n}(\mathbb{R})$ est le nombre des colonnes linéairement indépendantes.

Exemple 1.1.1 Soit la matrice :

$$A = \begin{pmatrix} 10 & 2 & -1 \\ 1 & -8 & -1 \\ 1 & 2 & 6 \end{pmatrix}$$

$$\alpha_1 \begin{pmatrix} 10 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ -8 \\ 2 \end{pmatrix} + \alpha_3 \begin{pmatrix} -1 \\ -1 \\ 6 \end{pmatrix} = 0 \Leftrightarrow \alpha_1 = \alpha_2 = \alpha_3 = 0$$

Donc $\text{rang}(A) = 3$.

Définition 1.1.7 Soit $A \in M_n(\mathbb{K})$ on appelle valeur propre de A le scalaire λ qui vérifie l'équation:

$$|A - \lambda I| = 0.$$

Autrement dit s'il existe un vecteur v non nul tel que

$$Av = \lambda v.$$

On appelle vecteur propre associé à la valeur propre λ le vecteur non nul v qui vérifie l'équation

$$(A - \lambda I)v = 0 \quad (I \text{ la matrice identité d'ordre } n).$$

L'ensemble des valeurs propres de A appelés le spectre de A et noté $Sp(A)$

$$Sp(A) = \{\lambda_i \in \mathbb{C}, i = 1, \dots, n\}.$$

Définition 1.1.8 Image et noyau d'une matrice $A \in M_{m,n}(\mathbb{R})$

$$\text{Im}(A) = \{y \in \mathbb{R}^m / y = Ax\}.$$

$$\text{Ker}(A) = \{x \in \mathbb{R}^n / Ax = 0\}.$$

1.1.1 Les normes vectorielles et les normes matricielles

Normes sur $\mathbb{R}^n$

On se place dans $\mathbb{R}^n$ (ensemble des vecteurs réels d'ordre n). On notera $\{e_1, \dots, e_n\}$ la base canonique de $\mathbb{R}^n$ le vecteur e_i a toutes ses composantes nulles sauf la i -ème qui vaut 1. On rappelle que $\mathbb{R}^n$ est un espace euclidien de dimension finie quand on le munit du produit scalaire

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i.$$

Définition 1.1.9 (Norme vectorielle) Soit X un espace vectoriel sur $\mathbb{K}$, une norme vectorielle sur X est une application noté $\|\cdot\|$ tel que :

$$\begin{aligned} \|\cdot\| &: X \rightarrow \mathbb{R}_+ \\ x &\rightarrow \|x\| \end{aligned}$$

vérifiant pour tout x, y dans X et α dans $\mathbb{K}$

- 1- $\|x\| = 0$ si et seulement si $x = 0$.
- 2- $\|\alpha x\| = |\alpha| \|x\|$ (homogénéité).
- 3- $\|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire).

Quelques normes vectorielles: Soit $x \in \mathbb{R}^n$ on a

$$\begin{aligned} \|x\|_1 &= \sum_{i=1}^n |x_i|. \\ \|x\|_2 &= \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}}. \\ \|x\|_\infty &= \max_{1 \leq i \leq n} \{|x_i|\}. \end{aligned}$$

Exemple 1.1.2 Soit le vecteur $x = (-1, 3, 5)^T$

on a

$$\|x\|_1 = 9, \quad \|x\|_2 = \sqrt{35}, \quad \|x\|_\infty = 5.$$

Définition 1.1.10 (Norme matricielle) Une norme matricielle est une application qui associe à une matrice A un scalaire noté $\|A\|$ vérifiant les propriétés suivantes:

Pour tout $A, B \in M_{m,n}(\mathbb{K})$ et $\lambda \in \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$).

- 1- $\|A\| > 0$ et $\|A\| = 0 \Leftrightarrow A = 0$.
- 2- $\|\lambda A\| = |\lambda| \|A\|$.
- 3- $\|B + A\| \leq \|A\| + \|B\|$.
- 4- $\|BA\| \leq \|A\| \|B\|$.

Quelques normes matricielles: Soit $A \in M_{m,n}(\mathbb{R})$

$$\begin{aligned} \|A\|_1 &= \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|. \\ \|A\|_2 &= \sqrt{\rho(AA^T)} \text{ avec } \rho(AA^T) = \max_{\lambda_i \in Sp(AA^T)} |\lambda_i|. \\ \|A\|_\infty &= \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|. \end{aligned}$$

Exemple 1.1.3 soit la matrice :

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix}$$

$$\|A\|_1 = 5, \quad \|A\|_2 = 12 + 2\sqrt{6}, \quad \|A\|_\infty = 6.$$

1.1.2 Conditionnement d'un système linéaire

On considère le système linéaire:

$$Ax = b,$$

où A est une matrice inversible, dans la pratique A et b peuvent être entachés à des erreurs.

Si b est perturbé tel que le système $Ax = b$ devient $A(x + \Delta x) = b + \Delta b$, le but est de mesurer la sensibilité de l'erreur relative commise sur x , i. e. $\delta x = \frac{\|\Delta x\|}{\|x\|}$ en fonction des données au problème, $\delta b = \frac{\|\Delta b\|}{\|b\|}$.

Conditionnement d'une matrice

Soit $\|\cdot\|$ une norme matricielle subordonnée, le conditionnement d'une matrice régulière A , associé à cette norme, est le nombre

$$\text{cond}(A) = \|A\| \|A^{-1}\|.$$

Nous notons

$$\text{cond}_p(A) = \|A\|_p \|A^{-1}\|_p.$$

Théorème 1.1.2 *Soit A une matrice inversible. Soient x et $x + \Delta x$ les solutions respectives de $Ax = b$ et $A(x + \Delta x) = b + \Delta b$, Si $b \neq 0$, alors*

$$\frac{\|\Delta x\|}{\|x\|} \leq \text{cond}(A) \frac{\|\Delta b\|}{\|b\|}.$$

Démonstration. : *Il suffit de soustraire les deux équations*

$$Ax = b \quad \text{et} \quad A(x + \Delta x) = b + \Delta b$$

d'où

$$\Delta x = A^{-1} \Delta b$$

alors

$$\|\Delta x\| \leq \|A^{-1}\| \|b\| \frac{\|\Delta b\|}{\|b\|} \leq \|A^{-1}\| \|A\| \|x\| \frac{\|\Delta b\|}{\|b\|}.$$

Donc

$$\frac{\|\Delta x\|}{\|x\|} \leq \text{cond}(A) \frac{\|\Delta b\|}{\|b\|}.$$

■

Théorème 1.1.3 *Soit A une matrice inversible. Soient x et $x + \Delta x$ les solutions respectives de*

$$Ax = b \quad \text{et} \quad (A + \Delta A)(x + \Delta x) = b$$

Si $b \neq 0$, alors

$$\frac{\|\Delta x\|}{\|x + \Delta x\|} \leq \text{cond}(A) \frac{\|\Delta A\|}{\|A\|}.$$

Démonstration. : Il suffit de soustraire les deux équations

$$Ax = b \text{ et } (A + \Delta A)(x + \Delta x) = b,$$

il vient

$$\Delta x = -A^{-1}\Delta A(x + \Delta x),$$

d'où

$$\|\Delta x\| \leq \|A^{-1}\| \|\Delta A\| \|x + \Delta x\|$$

donc

$$\frac{\|\Delta x\|}{\|x + \Delta x\|} \leq \|A^{-1}\| \|A\| \frac{\|\Delta A\|}{\|A\|} = \text{cond}(A) \frac{\|\Delta A\|}{\|A\|}.$$

Exemple 1.1.4 Soit le système $Ax = b$ où

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}.$$

Soit les deux systèmes perturbés

$$(A + \Delta A)(x + \Delta x) = b \text{ et } A(x + \Delta x) = b + \Delta b,$$

Avec

$$\Delta A = \begin{pmatrix} 0 & 0 & 0,1 & 0,2 \\ 0,08 & 0,04 & 0 & 0 \\ 0 & -0,02 & -0,11 & 0 \\ -0,01 & -0,01 & 0 & -0,02 \end{pmatrix}, \quad \Delta b = \begin{pmatrix} 0,01 \\ -0,01 \\ 0,01 \\ -0,01 \end{pmatrix}.$$

■

$$Ax = b \Leftrightarrow x = (1, 1, 1, 1)^T.$$

$$Au = b + \Delta b \Leftrightarrow u = \begin{pmatrix} 1,82 \\ -0,36 \\ 1,35 \\ 0,79 \end{pmatrix}, \Delta x = u - x = \begin{pmatrix} 0,82 \\ -1,36 \\ 0,35 \\ -0,21 \end{pmatrix}.$$

$$\text{cond}_\infty(A) = \|A\| \|A^{-1}\| = 33 \times 136 = 4488.$$

$$\frac{\|\Delta x\|_\infty}{\|x\|_\infty} = 1,36, \frac{\|\Delta b\|_\infty}{\|b\|_\infty} = 3,0303 \times 10^{-4}, \text{cond}(A)_\infty \frac{\|\Delta b\|_\infty}{\|b\|_\infty} = 1,36.$$

$$(A + \Delta A)v = b \Leftrightarrow v = \begin{pmatrix} -81 \\ 137 \\ -34 \\ 22 \end{pmatrix}, \Delta x = v - x = \begin{pmatrix} -82 \\ 136 \\ -35 \\ 21 \end{pmatrix}.$$

$$\frac{\|\Delta x\|_\infty}{\|x + \Delta x\|_\infty} = 0.9927, \frac{\|\Delta A\|}{\|A\|} = 0.0091, \text{cond}(A)_\infty \frac{\|\Delta A\|_\infty}{\|A\|_\infty} = 40,8408.$$

Système linéaire bien conditionné et mal conditionné

Définition 1.1.11 *Un système linéaire est mal conditionné si une petite variation des données entraîne une très grande variation des résultats. Un système linéaire qui n'est pas mal conditionné est dit bien conditionné.*

Si $\text{cond}(A)$ est petit (proche de 1), on dit que A est une matrice bien conditionnée dans ce cas le système $Ax = b$ est stable (bien conditionné) numériquement la solution x existe et unique, et une petite variation des données (A, b) à fournis une petite variation des résultats.

Si $\text{cond}(A) \gg 1$, on dit que A est une matrice mal conditionnée dans ce cas le système $Ax = b$ n'est pas stable numériquement, donc une petite variation des données (A, b) à fournis une grande variation des résultats.

Exemple 1.1.5 *Soit le système linéaire $Ax = b$*

où

$$A = \begin{pmatrix} 4.218613 & 6.327917 \\ 3.141592 & 4.71239 \end{pmatrix}, b = \begin{pmatrix} 10.546530 \\ 7.853982 \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

$$Ax = b \Leftrightarrow x_1 = x_2 = 1.$$

Faisons un petit changement sur A nous trouvons :

$$\begin{pmatrix} 4.218611 & 6.327917 \\ 3.141594 & 4.71239 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10.546530 \\ 7.853982 \end{pmatrix} \Leftrightarrow x_1 = -5, x_2 = 5.$$

$\text{cond}_\infty(A) = 7.14726 \times 10^6$ très grand, donc la matrice A est mal conditionnée et le système instable.

Exemple 1.1.6 (*Matrice de Hilbert*)

Une matrice de Hilbert est une matrice carrée de terme générale

$$H = (h_{i,j})_{1 \leq i, j \leq n} = \frac{1}{i+j-1}$$

$$H = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$$

$\text{cond}_\infty(H) = 747.9864 \gg 1$ très grand, donc la matrice H est mal conditionnée.

Exemple 1.1.7 (*Matrice de Vandermonde*)

Une matrice de Vandermonde est une matrice carrée prendre la forme

$$V = [v^0, v, v^2, \dots, v^{n-1}] \text{ avec } v \in \mathbb{R}^n$$

$$V = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$$

$\text{cond}_\infty(V) = 104 \gg 1$ très grand, donc la matrice V est mal conditionnée

1.2 Dérivation d'une forme quadratique

Définition 1.2.1 Une forme quadratique est un polynôme homogène de degré 2 avec un nombre quelconque des variables :

$$\begin{aligned} f &: \mathbb{R}^n \rightarrow \mathbb{R} \\ x &= (x_1, x_2, \dots, x_n) \rightarrow \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \end{aligned}$$

en note $A = (a_{ij}) \in M_{n,n}(\mathbb{R})$, f s'écrit

$$\begin{aligned} f &: \mathbb{R}^n \rightarrow \mathbb{R} \\ x &\mapsto x^T A x = \langle x, Ax \rangle = \langle Ax, x \rangle \end{aligned}$$

Définition 1.2.2 On appelle dérivée directionnelle de f en x dans la direction d et on note $Df(x, d)$ la limite, quand elle existe

$$Df(x, d) = \lim_{\varepsilon \rightarrow 0} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon}.$$

Le calcul de la dérivée d'une forme quadratique s'obtient simplement à l'aide des dérivées directionnelles :

On considère

$$f(x) = \langle Ax, x \rangle$$

donc

$$\begin{aligned} Df(x, d) &= \lim_{\varepsilon \rightarrow 0} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon} \\ Df(x, d) &= \lim_{\varepsilon \rightarrow 0} \frac{\langle A(x + \varepsilon d), x + \varepsilon d \rangle - \langle Ax, x \rangle}{\varepsilon} \\ Df(x, d) &= \lim_{\varepsilon \rightarrow 0} \frac{\langle Ax, x \rangle + \varepsilon \langle Ax, d \rangle + \varepsilon \langle Ad, x \rangle + \varepsilon^2 \langle Ad, d \rangle - \langle Ax, x \rangle}{\varepsilon} \\ Df(x, d) &= \lim_{\varepsilon \rightarrow 0} \langle Ax, d \rangle + \langle Ad, x \rangle + \varepsilon \langle Ad, d \rangle \\ Df(x, d) &= \langle Ax, d \rangle + \langle A^T d, x \rangle \\ Df(x, d) &= \langle (A + A^T)x, d \rangle \end{aligned}$$

donc

$$\forall x \in \mathbb{R}^n, A \in M_{nn}, \frac{\partial (x^T A x)}{\partial x} = (A + A^T)x$$

si A symétrique i. e. si $A = A^T$ on a

$$\frac{\partial(x^T Ax)}{\partial x} = 2Ax.$$

Chapitre 2

Problème de moindres carrés linéaires

2.1 Propriétés mathématiques du problèmes de moindres carrés

2.1.1 Formulation matricielle

On suppose que l'on a résoudre un système linéaire

$$Ax = b,$$

où $A \in M_{m,n}(\mathbb{R})$ et $b \in \mathbb{R}^m$ non nul, et on suppose que $m > n$ (le nombre d'équations est supérieur strictement au nombre d'inconnues). Dans la plupart des cas, ce système n'a pas de solution. On cherche alors une approximation de la solution qui réduise la différence $Ax - b$. Un des choix possible est de minimiser

$$\|Ax - b\|_2$$

Problème de moindres carrés

Définition 2.1.1 Soit $A \in M_{m,n}(\mathbb{R})$ et $b \in \mathbb{R}^m$ donnés. On appelle problème de moindres carrés le problème

$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$ i. e. trouvé x telle que $\|Ax - b\|_2^2$ soit la plus petite. On notera $\hat{x}$ la solution de ce problème.

2.1.2 Les équations normales

on pose $E(x) = \|Ax - b\|_2^2$, on a

$$\begin{aligned} E(x) &= (Ax - b)^T (Ax - b), \\ E(x) &= x^T A^T A x - 2b^T (Ax) + \|b\|_2^2. \end{aligned}$$

Ainsi,

$$\begin{aligned} \frac{\partial}{\partial x} E(x) &= \frac{\partial(x^T A^T A x)}{\partial x} - 2 \frac{\partial(b^T (Ax))}{\partial x} + \frac{\partial(b^T b)}{\partial x}, \\ \frac{\partial}{\partial x} E(x) &= 2A^T A x - 2A^T b. \end{aligned}$$

alors

$$\frac{\partial}{\partial x} E(x) (\hat{x}) = 0 \Leftrightarrow A^T A \hat{x} = A^T b. \quad (*)$$

$\hat{x}$ vérifiant un extrémum de $E(x)$

donc on montre que $\hat{x}$ vérifiant le minimum $E(x)$

L'existence de solution du problèmes du moindres carrés

Soient $A \in M_{m,n}(\mathbb{R})$ et $b \in \mathbb{R}^m$

Soit $\hat{x} \in \mathbb{R}^n$, on montre que:

$$\forall x \in \mathbb{R}^n, \|A\hat{x} - b\|_2^2 \leq \|Ax - b\|_2^2 \Leftrightarrow A^T A \hat{x} = A^T b.$$

On notera $\hat{b}$ la projection orthogonale de b sur ImA , on a donc par définition $\hat{b} \in ImA$ et $b - \hat{b} \in (ImA)^\perp$

$$1 - \|A\hat{x} - b\|_2^2 \leq \|Ax - b\|_2^2, \forall x \in \mathbb{R}^n \Leftrightarrow \|A\hat{x} - b\|_2^2 \leq \|z - b\|_2^2, \forall z \in ImA \Leftrightarrow A\hat{x} = \hat{b}$$

$$A\hat{x} = \hat{b} \Leftrightarrow A^T A \hat{x} = A^T \hat{b}$$

$$2 - A^T A \hat{x} = A^T \hat{b} \Leftrightarrow A^T (A\hat{x} - \hat{b}) = 0 \Leftrightarrow A\hat{x} - \hat{b} \in \ker A^T = (ImA)^\perp$$

$$\text{or } \hat{b} \in ImA \text{ donc } A\hat{x} - \hat{b} \in ImA$$

$$A\hat{x} - \hat{b} \in ImA \cap (ImA)^\perp \Rightarrow A\hat{x} - \hat{b} = 0 \Leftrightarrow A\hat{x} = \hat{b}$$

Montrons maintenant que

$$A^T A \hat{x} = A^T \hat{b} \Leftrightarrow A^T A \hat{x} = A^T b.$$

On a $b - \hat{b} \in (Im A)^\perp = \ker A^T$ donc $A^T(b - \hat{b}) = 0 \Leftrightarrow A^T b = A^T \hat{b}$

donc

$$\|A\hat{x} - b\|_2^2 \leq \|Ax - b\|_2^2, \forall x \in \mathbb{R}^n \Leftrightarrow A^T A\hat{x} = A^T b. \quad (**)$$

de (*) et (**), $\hat{x}$ vérifiant le minimum de $E(x)$.

Théorème 2.1.1 Soit $A \in M_{m,n}(\mathbb{R})$ avec $m > n$ et $b \in \mathbb{R}^m$. Une condition nécessaire et suffisante pour que $x \in \mathbb{R}^n$ réalise le minimum de $E(x) = \|Ax - b\|_2^2$ est que

$$A^T A\hat{x} = A^T b \quad (2.1.1)$$

Les équations (1) sont appelées équations normales. Ce système admet toujours au moins une solution. Si la matrice $A^T A$ est régulière, i. e. si $\text{rang}(A) = n$, alors la solution est unique.

Preuve. Nous montrons que les équations normales ont toujours au moins une solution, c'est-à-dire que $A^T b \in Im(A^T)$. Pour cela, commençons par remarquer que

$$Im A^T = Im (A^T A)$$

En effet,

$$Im(A^T A) = Ker(A^T A)^\perp = Ker(A)^\perp = Im(A^T)$$

Il suffit donc de vérifier que $A^T b \in Ker(A)^\perp$, or c'est immédiat, puisque, si $x \in Ker(A)^\perp$,

$$\langle A^T b, x \rangle = \langle b, Ax \rangle = 0 \quad \blacksquare$$

I'unicité de solution du problème de moindres carrés

On a $A \in M_{m,n}(\mathbb{R})$ avec $m > n$ $\text{rang} A = n$ et $b \in \mathbb{R}^m$

On suppose qu'il existe deux solutions de problème de moindres carrés x_1, x_2

Si x_1 est une solution de problème de moindres carrés, alors

$$\forall x \in \mathbb{R}^n \quad \|Ax_1 - b\|_2^2 \leq \|Ax - b\|_2^2.$$

En cas particulier $x = x_2$, d'où

$$\|Ax_1 - b\|_2^2 \leq \|Ax_2 - b\|_2^2 \quad (1)$$

Si x_2 est une solution de problème de moindres carrés, alors

$$\forall x \in \mathbb{R}^n \quad \|Ax_2 - b\|_2^2 \leq \|Ax - b\|_2^2.$$

En cas particulier $x = x_1$, d'où

$$\|Ax_2 - b\|_2^2 \leq \|Ax_1 - b\|_2^2 \quad (2)$$

de (1) et (2)

$$\|Ax_1 - b\|_2^2 = \|Ax_2 - b\|_2^2 \Leftrightarrow x_1 = x_2$$

2.2 Méthodes de résolution du problème de moindres carrés

2.2.1 Résolution du problème de moindres carrés par la méthode (QR)

La factorisation QR

Théorème 2.2.1 *La factorisation QR d'une matrice $A \in M_{m,n}(\mathbb{R})$, avec $\text{rang}(A) = n$ est une factorisation de la forme $A = QR$ où $Q \in M_{m,m}(\mathbb{R})$ est une matrice orthogonale ($Q^T Q = I$), et $R \in M_{m,n}(\mathbb{R})$ est une matrice triangulaire supérieure. Ce type de décomposition est souvent utilisé pour le calcul des solutions de systèmes linéaires non carrés, notamment pour déterminer la pseudo-inverse d'une matrice.*

Méthode de Householder Soient $A \in M_{m,n}(\mathbb{R})$, qui est représentée à l'aide de ses vecteurs colonnes selon : $A = [x_1, x_2, \dots, x_n]$

on pose : $A_1 = A$. Il faut d'abord effectuer une transformation de Householder qui transforme la première colonne x_1 de A en un vecteur $\|x_1\|e_1$ où $e_1 = (1, 0, \dots, 0)^T$.

Soit : $\alpha = \pm\|x_1\|$ (α doit de plus être du signe du premier élément de x), on pose

$$u = x_1 - \alpha e_1, \quad v = \frac{u}{\|u\|}.$$

la matrice de Householder est : $H_1 = I - 2vv^T$.

$$H_1 A_1 = \begin{pmatrix} \alpha_1 & * & * & * \\ 0 & & & \\ 0 & & A_2 & \\ 0 & & & \end{pmatrix}.$$

La matrice A_2 est de dimension inférieure d'une unité du point de vue nombre des lignes et des colonnes par rapport à A_1 . On peut répéter exactement l'opération sur cette matrice A_2 . Notons H'_2 la matrice de householder qui intervient. La raison de la notation H'_2 vient de sa dimension inférieure d'une unité par rapport à H_1 . Pour cette raison, on note H_2 déduite de H'_2 par :

$$H_2 = \begin{pmatrix} 1 & 0 \\ 0 & H'_2 \end{pmatrix}$$

est de manière générale :

$$H_k = \begin{pmatrix} I_{K-1} & 0 \\ 0 & H'_k \end{pmatrix}$$

où I_{K-1} est la matrice identité de dimension $K - 1$

Après t itérations, $t = \min(m - 1, n)$, on obtient les matrices de la factorisation QR :

$Q = H_1^T H_2^T \dots H_t^T$, $R = H_t H_{t-1} \dots H_1$. R est une matrice triangulaire supérieure.

Exemple 2.2.1 Soit la matrice A

$$A = \begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix}$$

$$Q_1 = \begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{7}{25} & -\frac{24}{25} \\ 0 & -\frac{24}{25} & -\frac{7}{25} \end{pmatrix}$$

$$Q = Q_1 Q_2 = \begin{pmatrix} \frac{6}{7} & \frac{69}{175} & \frac{-58}{175} \\ \frac{3}{7} & \frac{-158}{175} & \frac{6}{175} \\ -\frac{2}{7} & -\frac{6}{35} & -\frac{33}{35} \end{pmatrix}, \quad R = Q_2 Q_1 A = Q^T A = \begin{pmatrix} 14 & 21 & -14 \\ 0 & -175 & 70 \\ 0 & 0 & 35 \end{pmatrix}$$

Résolution du problème de moindres carrés par (QR) On a la factorisation QR précédente, ce qui permet d'écrire A sous la forme :

$$A = Q \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix}$$

où $\tilde{R} \in M_{n,n}(\mathbb{R})$ est triangulaire supérieure.

Notons que pour tout vecteur y de $\mathbb{R}^m$ on a

$$\|Q^T y\|_2^2 = (Q^T y)^T Q^T y = y^T Q Q^T y = y^T I y = \|y\|_2^2$$

puisque Q^T , comme Q , est orthogonale. On a donc, en particulier :

$$\|Ax - b\|_2^2 = \|Q^T(Ax - b)\|_2^2 = \|Q^T Ax - Q^T b\|_2^2$$

définissons $c \in \mathbb{R}^n$ et $d \in \mathbb{R}^{m-n}$ par

$$Q^T b = \begin{pmatrix} c \\ d \end{pmatrix}$$

alors

$$Q^T Ax - Q^T b = \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix} x - Q^T b = \begin{pmatrix} \tilde{R}x - c \\ d \end{pmatrix}$$

donc

$$\|Ax - b\|_2^2 = \left\| \begin{pmatrix} \tilde{R}x - c \\ d \end{pmatrix} \right\|_2^2 = \|\tilde{R}x - c\|_2^2 + \|d\|_2^2$$

Le vecteur $\hat{x}$ minimisant la norme de $\|Ax - b\|_2^2$ est donné par $\tilde{R}\hat{x} = c$

puisque $\|d\|_2^2$ est une constante qui ne joue aucun rôle dans la minimisation, le vecteur $\hat{x}$ est unique si $\tilde{R}$ est inversible, ce qui est le cas si $\text{rang}(A) = n$.

2.2.2 Résolution du problème de moindres carrés par décomposition de Cholesky

Décomposition de Cholesky

Théorème 2.2.2 *pour toute matrice $A \in M_{n,n}(\mathbb{R})$ symétrique et définie positive, il existe une matrice triangulaire inférieure unique $L \in M_{n,n}(\mathbb{R})$ dont tous les éléments diagonaux l_{ii} sont strictement positifs et qui satisfait*

$$A = LL^T, \text{ avec } l_{ik} = 0, \forall k > i, \text{ et } l_{ii} > 0, \forall i$$

les coefficients l_{ji} de L peuvent être calculés comme suit :

$$l_{11} = \sqrt{\alpha_{11}}, \quad l_{j1} = \frac{\alpha_{1j}}{l_{11}}$$

et pour $i = 2, \dots, n$

$$l_{ii} = (\alpha_{ii} - \sum_{k=1}^{i-1} l_{ik}^2)^{1/2} \text{ et } l_{ji} = \frac{1}{l_{ii}} (\alpha_{ij} - \sum_{k=1}^{i-1} l_{ik} l_{jk}), \quad j = i+1, \dots, n.$$

Cette factorisation $A = LL^T$ est parfois prendre la racine carrée de la matrice.

Démonstration. par récurrence: ■

Si $n = 1$, $A = \alpha > 0$. On peut écrire : $\alpha = l_{11} l_{11}^T$ d'où : $l_{11} = \sqrt{\alpha}$.

On suppose que le théorème est vrai jusqu'à l'ordre $n-1$ et on le démontre pour l'ordre n . La matrice $A \in M_{n,n}(\mathbb{K})$ peut être partitionnée selon :

$$A = \begin{pmatrix} A_{n-1} & b \\ b^T & \alpha_{nn} \end{pmatrix}$$

b est un vecteur de $\mathbb{K}^{n-1}$ et A_{n-1} est une matrice définie positive d'ordre $n-1$. Grâce à la propriété de récurrence, on sait qu'il existe une matrice unique L_{n-1} triangulaire inférieure d'ordre $n-1$ et à éléments diagonaux positifs telle que

$$A_{n-1} = L_{n-1} L_{n-1}^T, \quad l_{ik} = 0, \forall k > i \text{ et } l_{ii} > 0, \forall i$$

Supposons qu'il existe une matrice L telle que :

$$L = \begin{pmatrix} L_{n-1} & 0 \\ c^T & \alpha \end{pmatrix}$$

et on essaie de trouver le vecteur $c \in \mathbb{K}^{n-1}$ et le scalaire $\alpha > 0$ tel que :

$$\begin{pmatrix} L_{n-1} & 0 \\ c^T & \alpha \end{pmatrix} \begin{pmatrix} L_{n-1} & c \\ 0 & \alpha \end{pmatrix} = \begin{pmatrix} A_{n-1} & b \\ b^T & \alpha_{nn} \end{pmatrix} = A$$

donc :

$$L_{n-1}c = b \Rightarrow c = L_{n-1}^{-1}b \text{ (} L_{n-1} \text{ est inversible car } \det(A) = |\det(L_{n-1})|^2 \alpha^2 > 0 \text{)}$$

$$c^T c + \alpha^2 = \alpha_{nn} \Rightarrow \alpha = \sqrt{\alpha_{nn} - c^T c}$$

On supposant que tous les éléments l_{ij} sont connus jusqu'à la colonne $k-1$ incluse ($j \leq k-1$) les équations donnant les éléments suivante l_{kk} et l_{ik} ($i \geq k+1$)

$$\alpha_{kk} = |l_{k1}|^2 + |l_{k2}|^2 + \dots + |l_{kk}|^2 \text{ (} l_{kk} > 0 \text{)}$$

$$\alpha_{ik} = l_{i1} \overline{l_{k1}} + l_{i2} \overline{l_{k2}} + \dots + l_{ik} \overline{l_{kk}}$$

Les éléments de la matrice L sont égaux à:

$$l_{11} = \sqrt{\alpha_{11}}, \quad l_{j1} = \frac{\alpha_{1j}}{l_{11}}, \quad l_{ii} = (\alpha_{ii} - \sum_{k=1}^{i-1} l_{ik}^2)^{1/2}, \quad l_{ji} = \frac{1}{l_{ii}} (\alpha_{ij} - \sum_{k=1}^{i-1} l_{ik} l_{jk}), \quad j = i+1, \dots, n$$

Exemple 2.2.2 Soit la matrice symétrique définie positive :

$$A = \begin{pmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{pmatrix}$$

$$L = \begin{pmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{pmatrix}, \quad L^T = \begin{pmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

$$A = \begin{pmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{pmatrix} \begin{pmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

Résolution du problème de moindres carrés par la méthode de Cholesky Soit les équations normales $A^T A x = A^T b$

La matrice $A^T A$ soit mal conditionnée, et il est alors peu recommandé d'inverser cette matrice la résolution du système peut se faire en évitant d'inverser la matrice $A^T A$, ce qui augmente la robustesse de la résolution. Il faut procéder de la manière suivante :

Remarquons que la matrice $A^T A$ est une matrice symétrique définie positive, on effectue une décomposition de Cholesky donne : $LL^T = A^T A$, où la matrice L est triangulaire inférieure. L'équation normale $A^T A x = A^T b$ devient, alors $LL^T x = A^T b$.

posons $y = L^T x$ et $A^T b = B$, on obtient le système linéaire $Ly = B$ donc la solution de ce système est

$$y_i = \frac{1}{l_{ii}}(B_i - \sum_{j=1}^{i-1} l_{ij} y_j) \quad i = 1, \dots, n$$

y étant maintenant connu, on résout le système

$$L^T x = y$$

On pose $L^T = U$ d'où $Ux = y$ et par suite

$$x_i = \frac{1}{U_{ii}}(y_i - \sum_{j=i+1}^n U_{ij} x_j) \quad i = 1, \dots, n$$

2.2.3 Résolution du problème de moindres carrés par décomposition en valeurs singulières

Décomposition en valeurs singulières

On a vu que pour certaines matrices carrées on pouvait faire une décomposition en valeurs propres

$A = PDP^{-1}$, où D est une matrice diagonale de valeurs propres et P est une matrice inversible de vecteurs propres.

Mais pour effectuer cette transformation la matrice A doit être carrée et diagonalisable. L'idée de la décomposition en valeurs singulières est similaire à la décomposition en valeurs propres, mais fonctionne pour n'importe quelle matrice A de taille $m \times n$.

Définition 2.2.1 On appelle valeurs singulières d'une matrice réelle A d'ordre $m \times n$, les racines carrées des valeurs propres à la fois de la matrice carrée $A^T A$ et AA^T .

Rappelons que les matrices $A^T A$ et AA^T , sont semi-définies positives

Théorème 2.2.3 La décomposition en valeurs singulières (Singular Value Décomposition ou "SVD") d'une matrice rectangulaire $A \in M_{m, n}(\mathbb{R})$ de rang p , peut être exprimée comme le produit de trois matrices, une matrice orthogonale $U \in M_{m, m}(\mathbb{R})$, une matrice diagonale $\Sigma \in M_{m, n}(\mathbb{R})$ et la transposée d'une matrice orthogonale $V \in M_{n, n}(\mathbb{R})$

$$A = U\Sigma V^T.$$

où

$$\Sigma = (\Sigma_1, 0) \text{ si } n \geq m.$$

et

$$\Sigma = \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix} \text{ si } n \leq m.$$

telle que :

$$\Sigma_1 = \text{diag}(\mu_1, \mu_2, \dots, \mu_p), \mu_1 \geq \mu_2 \geq \dots \geq \mu_p, p = \min(m, n)$$

les μ_i sont les valeurs singulières de A et sont uniques.

Si les valeurs singulières sont toutes distinctes, les vecteurs singuliers à gauche u_i et à droite v_i sont uniques.

Preuve. Si $A = U\Sigma V^T$ alors $\|A\|_2 = \|\Sigma\|_2 = \mu_1$. Donc si la décomposition existe, μ_1 est unique. On va montrer l'existence de la décomposition.

Soit $\mu_1 = \|A\|_2$ et v_1 tel que $\|v_1\| = 1$, $\|u_1\| = 1$, $Av_1 = \mu_1 u_1$.

Soit $U_1 = (u_1, X)$ et $V_1 = (v_1, Y)$ deux matrices orthogonales d'ordre m et n

$$U_1^T A V_1 = \begin{pmatrix} u_1^T \\ X^T \end{pmatrix} A \begin{pmatrix} v_1 & Y \end{pmatrix} = \begin{pmatrix} u_1^T A v_1 & u_1^T A Y \\ X^T A v_1 & X^T A Y \end{pmatrix} = \begin{pmatrix} \mu_1 & w^T \\ 0 & B \end{pmatrix}$$

On va montrer que $w = 0$.

$$\|U_1^T A V_1\|_2 = \|A\|_2 = \mu_1$$

Soit $x = \begin{pmatrix} \mu_1 \\ w \end{pmatrix}$, $y = (U_1^T A V_1) x$. Alors

$$\begin{aligned} \|x\|_2^2 &= \mu_1^2 + w^T w \geq \mu_1^2 \\ \|y\|_2 &\leq \|U_1^T A V_1\|_2 \|x\|_2 = \mu_1 \|x\|_2 \end{aligned}$$

D'autre part,

$$y = \begin{pmatrix} \mu_1 & w^T \\ 0 & B \end{pmatrix} \begin{pmatrix} \mu_1 \\ w \end{pmatrix} = \begin{pmatrix} \mu_1^2 + w^T w \\ Bw \end{pmatrix}$$

donc

$$\|y\|_2 \geq \mu_1^2 + w^T w = \|x\|_2^2$$

d'où

$$\|x\|_2^2 \leq \mu_1^2 \text{ et } w = 0$$

Donc

$$U_1^T A V_1 = \begin{pmatrix} \mu_1 & 0 \\ 0 & B \end{pmatrix}$$

Nous allons maintenant faire une preuve par récurrence.

Si $A = u \in \mathbb{R}^{p \times 1}$ ou $A = v^T \in \mathbb{R}^{1 \times p}$ alors la décomposition existe.

On a $B \in \mathbb{R}^{(m-1) \times (n-1)}$. Supposons que $B = U_2 \Sigma_2 V_2^T$

$$U = U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}, \quad V = V_1 \begin{pmatrix} 1 & 0 \\ 0 & V_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \mu_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}$$

alors U et V sont des matrices orthogonales et

$$A = U \Sigma V^T$$

On a $\mu_2 = \|B\|_2 \leq \mu_1$

Il reste à montrer l'unicité des vecteurs singuliers. On l'admet ici

On a donc les égalités suivantes :

Si $n \leq m$, $p = n$:

$$\begin{aligned} Av_i &= \mu_i u_i, \quad i = 1, \dots, n, \\ A^T u_i &= \mu_i v_i, \quad i = 1, \dots, n, \\ A^T u_i &= 0, \quad i = n+1, \dots, m. \end{aligned}$$

Si $m \leq n$, $p = m$:

$$\begin{aligned} Av_i &= \mu_i u_i, \quad i = 1, \dots, m, \\ Av_i &= 0, \quad i = m+1, \dots, n, \\ A^T u_i &= \mu_i v_i, \quad i = 1, \dots, m. \end{aligned}$$

Ces égalités permettent de faire le lien avec les valeurs propres des matrices symétriques $A^T A$ et AA^T . ■

Calcul de la SVD

on a

$$A^T A = (U \Sigma V^T)^T (U \Sigma V^T) = V \Sigma^2 V^T = V \Sigma^2 V^{-1} \quad (a)$$

d'autre part,

$$A^T A = P D P^{-1} \quad (b)$$

Les colonnes de P est la base des vecteurs propres de $A^T A$ et D est une matrice diagonale contient les valeurs propres de $A^T A$

de (a) et (b)

$$\begin{aligned} \Sigma^2 &= D \Rightarrow \mu_i^2 = \lambda_i \Rightarrow \mu_i = \sqrt{\lambda_i} \Leftrightarrow \mu_i \geq 0 \\ \Sigma &= \text{diag}(\mu_i) \Leftrightarrow \Sigma^{-1} = \text{diag}(1/\mu_i) \Rightarrow \mu_i > 0 \\ V &= P \end{aligned}$$

$V = [v_1, \dots, v_i, \dots, v_r]$ tels que v_i les vecteurs propres de la matrice $A^T A$

on a

$$A = U \Sigma V^T$$

d'où

$$AV = U\Sigma V^T V \Leftrightarrow AV = U\Sigma$$

Donc

$$Av_i = \mu_i u_i$$

les u_i les colonnes de matrice U

Exemple 2.2.3 Soit la matrice

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix},$$

$$U = \begin{pmatrix} -0,7071 & 0 & -0,7071 \\ 0 & 1 & 0 \\ -0,7071 & 0 & 0,7071 \end{pmatrix}, V = \begin{pmatrix} -0,7071 & -0,7071 \\ -0,7071 & 0,7071 \end{pmatrix}, \Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Résolution du problème de moindres carrés par (DVS) La solution du problème de moindres carrés x donnée par $x = (A^T A)^{-1} A^T b$

pour $A = U\Sigma V^T$ on a

$$\begin{aligned} x &= ((U\Sigma V^T)^T (U\Sigma V^T))^{-1} (U\Sigma V^T)^T b \\ x &= (V\Sigma^2 V^T)^{-1} (U\Sigma V^T)^T b && \text{puisque } U^T U = I \\ x &= (V\Sigma^{-2} V^T) (V\Sigma U^T) b && \text{puisque } V^T V = I \\ x &= V\Sigma^{-1} U^T b && \Sigma^{-1} = \text{diag}\left(\frac{1}{\mu_i}\right) \\ x &= \sum_{i=1}^r \frac{u_i^T b}{\mu_i} v_i \end{aligned}$$

où $r = \text{rang}(A)$, et u_i et v_i sont respectivement le i -ème colonne de U et V dans le problème de moindres carrés :

$$\|Ax - b\|_2^2 = \|U\Sigma V^T x - b\|_2^2 = \|\Sigma V^T x - U^T b\|_2^2$$

puisque U est orthogonale. Notons $w = U^T b$ ($\|w\|_2 = \|b\|_2$), et prenons comme nouvelle inconnue $y = V^T x$ (avec encore $\|y\|_2 = \|x\|_2$), puisque V est orthogonale, ce qui sera important pour calculer la solution de norme minimale. Comme Σ est diagonale, ce problème est découplé, et se résout composante

par composante dans les bases $(u_1, \dots, u_m)$ et $(v_1, \dots, v_n)$.

Nous avons donc :

$$\|Ax - b\|_2^2 = \|\Sigma y - w\|_2^2 = \sum_{i=1}^p |\mu_i y_i - w_i|^2$$

On obtient donc toutes les solutions du problème (2.1.1) en posant

$$y_i = \begin{cases} \frac{w_i}{\mu_i} & \text{pour } i = 1, \dots, p. \\ \text{quelconque} & \text{pour } i = p + 1, \dots, n. \end{cases}$$

Chapitre 3

Résultats numériques

Dans ce chapitre, on donne des résultats numériques issus de cet algorithme et qui implémentés numériquement sous MATLAB

3.1 Exemple

Soit le système mal conditionné $Ax = b$

où

$$A = \begin{pmatrix} 1 & 274 & 2450 \\ 1 & 180 & 3254 \\ 1 & 375 & 3802 \\ 1 & 205 & 2838 \\ 1 & 86 & 2347 \end{pmatrix}, \quad b = \begin{pmatrix} 162 \\ 120 \\ 223 \\ 131 \\ 67 \end{pmatrix}.$$

$$\text{cond}(A) = 1.7576 \times 10^4$$

$$\text{rang}(A) = 3$$

Résultat d'algorithme des équations normales

$$A^T A = \begin{pmatrix} 5 & 1120 & 14691 \\ 1120 & 297522 & 3466402 \\ 14691 & 3466402 & 44608873 \end{pmatrix}$$

$$\text{rang}(A^T A) = 3$$

donc $A^T A$ est une matrice inversible

$$x = (A^T A)^{-1} A^T b = (7.0325, 0.5044, 0.0070)^T$$

$$G = Ax = (162.4043, 120.6153, 222.8193, 130.3140, 66.8471)^T$$

$$E = |G - b| = (0.4043, 0.6153, 0.1807, 0.6860, 0.1529)^T$$

Résultat d'algorithme de QR

On a

$$A = QR$$

$$Q = \begin{pmatrix} -0.4472 & 0.2315 & 0.7647 & -0.2782 & -0.2902 \\ -0.4472 & -0.2037 & -0.5443 & -0.3221 & -0.5988 \\ -0.4472 & 0.6992 & -0.3338 & -0.0761 & 0.4404 \\ -0.4472 & -0.0880 & 0.0324 & 0.8787 & -0.1382 \\ -0.4472 & 0.0809 & -0.6390 & -0.2023 & 0.5867 \end{pmatrix}$$

$$R = 10^3 \begin{pmatrix} -0.0022 & -0.5009 & -6.5700 \\ 0 & 0.2160 & 0.8132 \\ 0 & 0 & -0.8846 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

On a

$$\tilde{R}x = c$$

$$\tilde{R} = 10^3 \begin{pmatrix} -0.0022 & -0.5009 & -6.5700 \\ 0 & 0.2160 & 0.8132 \\ 0 & 0 & -0.8846 \end{pmatrix}$$

$$c = (-314.3912, 114.6376, -6.1934)^T$$

$$\text{rang}\tilde{R} = 3$$

donc

$$x = (\tilde{R})^{-1}c = (7.0325, 0.5044, 0.0070)^T$$

$$G = Ax = (162.4043, 120.6153, 222.8193, 130.3140, 66.8471)^T$$

$$E = |G - b| = (0.4043, 0.6153, 0.1807, 0.6860, 0.1529)^T$$

Résultat d'algorithme de Cholesky

On a

$$M = A^T A = \begin{pmatrix} 5 & 1120 & 14691 \\ 1120 & 297522 & 3466402 \\ 14691 & 3466402 & 44608873 \end{pmatrix}$$

$$M = LL^T = 10^3 \begin{pmatrix} 0.0022 & 0 & 0 \\ 0.5009 & 0.2160 & 0 \\ 6.5700 & 0.8132 & 0.8846 \end{pmatrix} 10^3 \begin{pmatrix} 0.0022 & 0.5009 & 6.5700 \\ 0 & 0.2160 & 0.8132 \\ 0 & 0 & 0.8846 \end{pmatrix}$$

$$B = A^T b = (703, 182230, 2164253)^T$$

$$y = (L^T)^{-1}B = (314.3912, 114.6376, 6.1934)^T$$

$$x = L^{-1}y = (7.0325, 0.5044, 0.0070)^T$$

Résultat d'algorithme décomposition en valeurs singulières

$$A = USV^T$$

$$U = \begin{pmatrix} 0.3678 & -0.4976 & -0.6749 & -0.2782 & -0.2902 \\ 0.4864 & 0.4351 & 0.3345 & -0.3221 & -0.5988 \\ 0.5702 & -0.4730 & 0.5014 & -0.0761 & 0.4404 \\ 0.4247 & 0.0934 & -0.1403 & 0.8787 & -0.1382 \\ 0.3503 & 0.5750 & -0.4019 & -0.2023 & 0.5867 \end{pmatrix}$$

$$S = 10^3 \begin{pmatrix} 6.6991 & 0 & 0 \\ 0 & 0.1673 & 0 \\ 0 & 0 & 0.0004 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} 6.8834 & 0.0053 & -0.0027 \\ 0.0053 & 0 & 0 \\ -0.0027 & 0 & 0 \end{pmatrix}$$

$$S^{-1} = \begin{pmatrix} 0.0001 & 0 & 0 & 0 & 0 \\ 0 & 0.006 & 0 & 0 & 0 \\ 0 & 0 & 2.6236 & 0 & . \end{pmatrix}$$

$$y = S^{-1}U^T b = (0.0484, -0.4968, -7.0329)^T$$

$$x = Vy = (7.0325, 0.5044, 0.0070)^T$$

$$G = Ax = (162.4043, 120.6153, 222.8193, 130.3140, 66.8471)^T$$

$$E = |G - b| = (0.4043, 0.6153, 0.1807, 0.6860, 0.1529)^T$$

Conclusion générale

Nous avons présenté dans ce mémoire une synthèse sur le problème de moindres carrés tel qu'on parle sur quelques propriétés mathématiques du problème de moindres carrés.

Nous avons ainsi exposé quelques méthodes de résolution du problème de moindres carrés (QR, Cholesky, DVS) et on remarque que décomposition en valeurs singulières fournit la solution la plus claire du problème des moindres carrés et meilleure méthode puisque la propriété de A est de *rang* $p < n$, est point fort de DVS contre des méthodes QR et Cholesky qui l'hypothèse de A est de *rang* n .

Ce travail est consacré aux résolutions les équations de la physique par la méthode de moindres carrés si pour ça on prendre des système mal conditionnés d'ordre (m, n) qui sont résultat de problème physiques et résoudres par la méthode de moindres carrés dans le part pratique

Bibliographie

- [1] A.Quarterni, R.Sacco, F.Saleri, Méthodes numériques pour le calcul scientifique, Springer, Paris, 2000
- [2] Copyright 2015 EDF R&D - Document diffusé sous licence GNU FDL
<http://www.gnu.org/copyleft/fdl.html>.
- [3] G. H. GOLUB, C. REINSCH "Singular value decomposition and least squares solutions" in "Handbook for Automatic Computation - Linear Algebra, Vol. 2" J. H. WILKINSON, C. REINSCH Editors _ SPINGER VERLAG (1971).
- [4] Jocelyne Erhel, Nabil Nassif, Bernard Philippe, "Calcul matriciel et systèmes linéaires" Beyrouth, Liban,
- [5] Michel KERN Problèmes inverses INRIA ,ROCQUENCOURT, BP105, 78153 LE CHESNAY, Michel.Kern@inria.fr2002/2003
- [6] Pierre Corriou Méthodes numériques et optimisation, Lavoisier, Paris, 2010
- [7] Tony Bourdier ESIAL- Mathématiques Numériques 2007 – 2008-
- [8] UTC, Équipe de Mathématiques Appliquées, Résolution des problèmes de moindres carrés, MT09-Analyse numérique élémentaire, Juin 2007

Résumé

Dans ce mémoire, nous présentons une étude théorique et numérique concernant une résolution des problèmes de moindres carrés.

Finalement, nous donnons des résultats numériques des quelques exemples assimilés dans équations physiques et résoudre par la méthode de moindres carrés est faite en Matlab et on termine ce mémoire par une conclusion.

Mots clés : moindres carrés, fonction d'approximation, problème d'approximation.

Abstract

In this memory, we present a theoretical and numerical study a resolution of least squares problems

Finally we give numerical results for some examples assimilated into physical equations and solved by least squares method is made in matlab and we ended this memory with a conclusion.

Key words: least squares, approximation functions, approximation problems