RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ DR.YAHIA FARES DE MEDEA
FACULTÉ DES SCIENCES
DÉPARTEMENT DE MATHÉMATIQUES ET INFORMATIQUE



Mémoire de fin d'étude

En vue de l'obtention du

DIPLÔME DE MASTER EN MATHÉMATIQUE
Option : Recherche Opérationnelle

Réalisé par

BOUDEKKES Bouchra et LARBI Ikram

Thème :

Factorisation des matrices et applications

Présenté le 22 Juin 2022 devant le jury composé de :

Présidente	M^{me}	AÏT AMRANE Nacima Rosa	U.Y.F.M.
Examinatrice	M^{me}	KOUIDER AIAD Soumia	U.Y.F.M.
Encadreur	M^{r}	HAMZA ABDELHALIM Younes	U.Y.F.M.

Année universitaire : 2021-2022

Remerciement

Toute notre gratitude et tous nos remerciements les plus sincères à ALLAH le tout puissant pour nous avoir guidés et assistés durant tous nos parcours d'études.

Nous tenons à remercier chaleureusement notre encadreur Monsieur HAMZA ABDELHALIM Younes pour le temps qu'il a consacré pour nous avoir apportés les outils méthodologiques indispensables à la conduite de ce mémoire et qui nous a offert la possibilité de réaliser ce travail dans de bonnes conditions, et aussi pour ses bénéfiques conseils et explications, nous avons eu l'honneur de travailler sous sa direction.

A Madame AÏT AMRANE Nacima Rosa d'avoir accepté de présider notre jury pour évaluer notre travail. Vos qualités scientifiques et votre intérêt pour la recherche sont pour nous une source de motivation supplémentaire pour la suite de nos carrières. Recevez ici, toute notre gratitude et toute notre sincère reconnaissance pour votre disponibilité.

Les plus grandes leçons ne sont pas tirées d'un livre, mais d'un enseignant tel que vous Madame KOUIDER AIAD Soumia. Nous tenons à vous remercier pour votre enseignement et votre soutien tout au long des années d'études.

Nous voudrions aussi adresser nos remerciements les plus sincères vivement nos famille, nos amis, et tous ce qui ont aidé dans la réalisation de ce mémoire .

Dédicace

Avec l'expression de ma reconnaissance, je dédie de ce modeste travail à mon exemple dans cette vie, à ma chère mère "Lila".

Quoi que je fasse ou que je dise, je ne saurai point vous remercier comme il se doit.Votre affectation me couvre , votre bienveillance me guide et votre présence à mes côtes a toujours été ma source de force pour affronter les différents obstacles.

Á mon cher père "Mohamed", Vous avez toujours été à mes côtés pour me soutenir et m'encourager. Votre affectation et la confiance qui vous m'a accordé.

Á mon très cher grand-père "Moustapha qui m'a fait de moi une femme et à la femme qui est ma deuxième maman ,qui m'a protégé par ses prières,ma chère grand-mère "Fatiha".

Á Mes chers Oncles et mes chères tantes :"Brahem,Ryad,Mohamed,Chakir,Amina,Ratiba".

Á mes mes cousins et mes cousines.

Son oublier mon binôme "Ikram", pour son soutien moral, sa patience et sa compréhension tout au long de ce travail , mon intime amie et ma sœur avec qui j'ai partagé de beaux moments, que dieu te protège.

Á tous mes chères amies

Á tous ceux qui, même un mot, m'ont donné la force de continuer. . .

★★★ *Bouchra*★★★

Dédicace

C'est avec de profonds sentiments, de respect et d'amour, que je dédie ce travail.

Á "mes parents " que j'aime énormément, pour leur présence et leur soutien, pour leur dévouement mais surtout pour leur confiance en moi.

Á la source de laquelle j'ai toujours puisé soutient, courage et persévérance mon très cher père.
Tu as toujours été pour moi un exemple de la parfaite mère, un exemple d'honnêteté, de sérieux et de persévérance. chère et adorable mère.

Á ma chère grand-mère, que Dieu la protège et prolonge sa vie ,et Á mon grand-père, que Dieu lui fasse miséricorde.

Á ma chère tante "Nassima" .
Á mon cher frère "Hicham" .
Á ma sœur "Inés " et ma petite sœur "Ranim" qui me sont très chères,et mes chers frères "Youceff" et "Amin ".

Á mon cher fiancé "Moustapha " , et à toute sa famille.

Á mon binôme et mon intime amie et ma sœur "Bouchra " avec qui j'ai partagé de beaux moments, que dieu te protège .

Á mes chères amies
Á toute ma famille pour leur encouragements.
Á tous ceux que j'aime et ceux qui m'ont aidé.
Á tous ceux qui me sont chers et dont je n'ai pas cité leurs noms.

★★★ *Ikram*★★★

Résumé

Ce mémoire de master a pour but de faire une étude de certaines façons de factoriser les matrices. Nous avons mis les rappels nécessaires à notre étude, par une façon pas très répandue mais très utiles, de multiplier les matrices, méthode dite colonne par ligne.
Puis nous avons décrire les projections orthogonales puis déduits le procédé d'orthogonalisation de Gram-Schmidt.

Ensuite nous avons déduit le théorème de la décomposition spectrale d'une matrice symétrique à partir du théorème fondamental. Après nous avons déduit la décomposition en valeurs singuliers on a utilisé ce résultat pour définir le concept de pseudo-inverse.
Finalement nous avons appliqué la factorisation spectrale à l'étude des espaces d'ajustements aux dimensions supérieurs.

Mots-Clés : Factorisation des matrices, orthogonalisation, la projection, théorème de la décomposition spectrale, la décomposition en valeurs singuliers (SVD)

Abstract

The aims of This master's are to study certain ways of factoring matrices. We have put the reminders necessary for our study, by a way not very widespread but very useful, to multiply the matrices, method called column by line.

Then we described the orthogonal projections then deduced the Gram-Schmidt orthogonalization process.

Then we deduced the theorem of the spectral decomposition of a symmetric matrix from the fundamental theorem. After we deduced the decomposition into values singular ones we used this result to define the concept of pseudo-inverse.

Finally we applied the spectral factorization to the study of the spaces of adjustments to higher dimensions.

Keywords :Matrix factorization, orthogonalization, projection, spectral decomposition theorem, singular value decomposition (SVD)

Notation

$\mathbb{R}$	Corps des nombres réels.
$\mathbb{N}$	Corps des nombres naturels.
$M_n(\mathbb{R})$	L'ensemble des matrices carrées d'ordre n.
E	L'espace vectoriel.
$\mathbb{R}^n$	L'espace vectoriel des vecteurs à composantes réels, de dimension n.
$\{e_1, e_2, e_3\}$	La base canonique de $\mathbb{R}^3$.
I_n	La matrice identité d'ordre n .
$A^\top$	La matrice transposée de la matrice A .
A^{-1}	La matrice inverse de la matrice A .
$detA$	Le déterminant de la matrice carrée A .
$r(A)$	Le rang de la matrice A.
$ker(A)$	Le noyau de la matrice A.
$Im(A)$	L'image de la matrice A .
$tr(A)$	La trace de la matrice A .
P	La projection.
$U^\perp$	Le sous-espace vectoriel des vecteurs orthogonaux à U .
$(u\|v)$	Le produit scalaire de deux vecteurs de $\mathbb{R}^n$.
col_i^A	Les colonnes de la matrice A .
L_i^A	Les lignes de la matrice A .
λ_i	Les valeurs propres de la matrice A .
δ_{ij}	Le symbole de Kronecker .
$\oplus$	Somme directe .

TABLE DES MATIÈRES

Remerciement i

Dédicace ii

Dédicace iii

Résumé iv

Abstract v

Notation vi

Table des matières vii

Table des figures viii

Introduction 1

1 Rappels 2
- 1.1 Multiplication matricielle . . . 2
  - 1.1.1 Multiplication par la Méthode ligne colonne . . . 2
  - 1.1.2 Multiplication par la méthode colonne . . . 3
  - 1.1.3 Multiplication par la méthode ligne . . . 4
  - 1.1.4 Multiplication colonnes par lignes . . . 6
- 1.2 Somme directe et espaces supplémentaires . . . 7
  - 1.2.1 Unicité de la décomposition et propriétés de l'intersection dans le cas de la somme directe . . . 7
  - 1.2.2 Bases d'une somme supplémentaire . . . 8
  - 1.2.3 Espaces orthogonaux . . . 8
  - 1.2.4 Projection sur un sous-espace vectoriel dans le cas d'une somme directe . 9
- 1.3 Détermination des valeurs et vecteurs propres d'une matrice . . . 10

2 Espace vectoriel et projection 11
- 2.1 Expression d'une application linéaire . . . 11
- 2.2 Sous-espace et projection . . . 12

2.3 Matrices de projection . . . 14
2.4 Projection sur des droites vectorielles et des plans (Hyperplans) vectoriel. . . . 17
2.5 Procédé d'orthogonalisation de Gram-Schmidt . . . 18

3 Valeurs propres et décomposition spectrale 21
3.1 Valeurs propres et vecteurs propres . . . 21
3.2 Décomposition Spectrale . . . 25
3.3 Décomposition d'une matrice symétrique . . . 26
3.4 Diagonalisation des matrices symétriques . . . 27
3.5 Inverse et Puissance . . . 28

4 Valeurs singulières et décomposition en valeur singulier (SVD) 31
4.1 Valeurs singuliers et vecteurs singuliers . . . 31
4.2 Espace colonne et espace ligne d'une matrice . . . 36
4.3 Représentation matricielle . . . 37
4.4 Pseudo-inverse . . . 38
4.5 Pseudo inverse d'un vecteurs . . . 42

5 Espaces d'ajustements 43
5.1 Droite d'ajustement . . . 43
5.2 Espace d'ajustement . . . 44
5.2.1 Extremums d'une forme quadratique . . . 45
5.2.2 Espace d'ajustements hiérarchisés . . . 47

Conclusion 51

TABLE DES FIGURES

2.1 La projection orthogonale de P sur U 13
2.2 La projection orthogonale de P sur un hyperplan vectoriel Π de $\mathbb{R}^n$ 18
2.3 Orthonormalisation de Gram-Schmidt 19

INTRODUCTION

Dans ce modeste mémoire nous avons étudié certaines façons de factoriser les matrices dans le but de voir comment ces factorisations permettent de définir de nouveaux concepts algébriques et comment elles permettent de déduire des résultats déjà connu de manière plus simple.

Dans le 1^{ier} chapitre nous avons mis les rappels nécessaires à notre étude, entre une façon pas très répandue mais très utiles, de multiplier les matrices, méthode dite colonne par ligne..

Dans le $2^{ième}$ chapitre nous avons, à l'aide de la multiplication colonne par ligne, montré comment on peut décomposer une matrice comme somme de produit colonne ligne dans une base orthonormée de l'espace de départ. Puis nous avons appliqué le résultat pour décrire les projections orthogonales puis déduits le procédé d'orthogonalisation de Gram-Schmidt..

Dans le $3^{ième}$ chapitre nous avons déduit le théorème de la décomposition spectrale d'une matrice symétrique à partir de théorème fondamental, puis déduits certaines opérations sur les matrices telles que la racine carré, et les puissances..

Dans le $4^{ième}$ chapitre nous avons défini ce qui sont les valeurs singulières d'une matrice ainsi que les vecteurs singuliers, puis on a montré comment une matrice quelconque se factorise dans les bases singulières , on a utilisé ce résultat pour définir le concept de pseudo-inverse..

Dans le $5^{ième}$ chapitre nous avons appliqué la factorisation spectrale à l'étude des espaces d'ajustements aux dimensions supérieurs.

Pour des notions plus précises et détaillées, consulter les références [1], [2], [3], [4],[5],[6], [7],[8], [9],[10], [11].

CHAPITRE

1

RAPPELS

Dans ce chapitre, on introduit la multiplication matricielle par différentes méthodes. On parlera de la multiplication par la méthode colonne par ligne , la méthode colonne et la méthode ligne.

On définit la somme directe de espaces supplémentaires. Puis On présente l'unicité de la décomposition et on décrira les bases d'une somme supplémentaire. On parlera aussi d'espaces orthogonaux et la projection sur un sous-espaces vectoriel.

1.1 Multiplication matricielle

1.1.1 Multiplication par la Méthode ligne colonne

En règle générale :

Soit $\begin{pmatrix} a_1 & a_2 & \cdots & a_p \end{pmatrix}$ et $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}$ deux vecteurs d'ordre p.

Alors

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_p \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_pb_p.$$

Soit la matrice $A = (a_{ij})$ une matrice d'ordre $n \times p$ et soit $B = (b_{ij})$ une matrice d'ordre $p \times m$. Alors

1. La multiplication matricielle AB est possible car le nombre p de colonnes de la matrice A est égale aux nombres de lignes p de la matrice B.

2. Le résultat de la multiplication AB est une matrice $C = (c_{ij})$ d'ordre $n \times m$ et dont les composantes c_{ij} sont définies par la multiplication de la $i^{éme}$ ligne de A par la $j^{éme}$ colonne de B.
Autrement dit :

$$c_{ij} = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{ip} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix}$$

$$= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj}$$

$$= \sum_{k=1}^{p} a_{ik}b_{kj}.$$

1.1.2 Multiplication par la méthode colonne

Si A est une matrice d'ordre $n \times p$ et si B est une matrice d'ordre $p \times m$, et si on désigne par $col_1^B, col_2^B, \cdots, col_m^B$ les colonnes de la matrice B, alors le produit AB sera une matrice à m colonnes et on a :

$$AB = A \begin{pmatrix} col_1^B & col_2^B & \cdots & col_m^B \end{pmatrix}$$

$$= \begin{pmatrix} A.col_1^B & A.col_2^B & \cdots & A.col_m^B \end{pmatrix}.$$

Autrement dit :

La multiplication d'un vecteur colonne par une matrice est une combinaison linéaire des vecteurs colonnes de la matrice.

Par induction, si on désigne les vecteurs colonnes de la matrices A par $col_1^A, col_2^A, \cdots, col_p^A$, et si on désigne par $col_j^B = \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix}$.

Alors on obtient :

$$A.col_j^B = \begin{pmatrix} col_1^A & col_2^A & \cdots & col_p^A \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix}$$

$$= b_{1j}col_1^A + b_{2j}col_2^A + \cdots + b_{pj}col_p^A.$$

1.1.3 Multiplication par la méthode ligne

1. Si A est une matrice d'ordre $n \times p$ dont les n lignes sont désignées par $L_1^A, L_2^A, \cdots, L_n^A$ et$\begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix}$ vecteur ligne .
   Alors symboliquement on aura :

$$A = \begin{pmatrix} L_1^A \\ L_2^A \\ \vdots \\ L_n^A \end{pmatrix}$$

   Alors

$$\begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix} \begin{pmatrix} L_1^A \\ L_2^A \\ \vdots \\ L_n^A \end{pmatrix} = \beta_1 L_1^A + \beta_2 L_2^A + \cdots + \beta_n L_n^A.$$

2. Si B une matrice d'ordre $m \times n$ et $L_1^B, L_2^B, \cdots, L_m^B$ sont les lignes de la matrice B alors la multiplication de B par une matrice A d'ordre $n \times p$ est donner comme suite :

$$BA = \begin{pmatrix} L_1^B \\ L_2^B \\ \vdots \\ L_m^B \end{pmatrix} A = \begin{pmatrix} L_1^B A \\ L_2^B A \\ \vdots \\ L_m^B A \end{pmatrix}.$$

   Avec

$$L_k^B A = \begin{pmatrix} b_{k1} & b_{k2} & \cdots & b_{km} \end{pmatrix} \begin{pmatrix} L_1^A \\ L_2^A \\ \vdots \\ L_m^A \end{pmatrix}$$

$$= b_{k1} L_1^A + b_{k2} L_2^A + \cdots + b_{km} L_m^A.$$

Exemple 1

Si on veut ajouter la $1^{ière}$ *ligne de la matrice* $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 2 & 3 \\ 2 & -1 & 0 \end{pmatrix}$ *à sa* $2^{iéme}$ *ligne sans changer la* $1^{ière}$ *et la* $3^{ième}$ *ligne on multiplie à gauche cette matrice par la matrice* $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, *en effet*

on a :

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} L_1^A \\ L_2^A \\ L_3^A \end{pmatrix} = \begin{pmatrix} 1L_1^A + 0L_2^A + 0L_3^A \\ 1L_1^A + 1L_2^A + 0L_3^A \\ 0L_1^A + 0L_2^A + 1L_3^A \end{pmatrix}$$

$$= \begin{pmatrix} L_1^A \\ L_1^A + L_2^A \\ L_3^A \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 1 \\ 0 & 4 & 4 \\ 2 & -1 & 0 \end{pmatrix}.$$

Si on veut multiplier la première ligne de la matrice A par (-2) *sans changer les autres lignes. Alors on multiplie à gauche la matrice A par la matrice élémentaire* $\begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, *en effet on a :*

$$\begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} L_1^A \\ L_2^A \\ L_3^A \end{pmatrix} = \begin{pmatrix} -2L_1^A + 0L_2^A + 0L_3^A \\ 0L_1^A + 1L_2^A + 0L_3^A \\ 0L_1^A + 0L_2^A + 1L_3^A \end{pmatrix}$$

$$= \begin{pmatrix} -2L_1^A \\ L_2^A \\ L_3^A \end{pmatrix}$$

$$= \begin{pmatrix} -2 & -4 & -2 \\ -1 & 2 & 3 \\ 2 & -1 & 0 \end{pmatrix}.$$

Si on veut retrancher 2 *fois la première ligne de la troisième ligne de la matrice A sans changer la* $2^{iéme}$ *ligne et la* $1^{ième}$ *ligne de la matrice A on la multiplie a gauche par la matrice* $E_{31}(-2)$ $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$, *en effet on a :*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} L_1^A \\ L_2^A \\ L_3^A \end{pmatrix} = \begin{pmatrix} 1L_1^A + 0L_2^A + 0L_3^A \\ 0L_1^A + 1L_2^A + 0L_3^A \\ -2L_1^A + 0L_2^A + 1L_3^A \end{pmatrix}$$

$$= \begin{pmatrix} L_1^A \\ L_2^A \\ -2L_1^A + L_3^A \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 1 \\ -1 & 2 & 3 \\ 0 & -5 & -2 \end{pmatrix}.$$

Si on veut permuter la $1^{ière}$ et la $2^{iéme}$ ligne de la matrice A sans changer la $3^{ième}$ ligne on multiplie la matrice A à gauche par la matrice $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, *en effet on a :*

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} L_1^A \\ L_2^A \\ L_3^A \end{pmatrix} = \begin{pmatrix} 0L_1^A + 1L_2^A + 0L_3^A \\ 1L_1^A + 0L_2^A + 0L_3^A \\ 0L_1^A + 0L_2^A + 1L_3^A \end{pmatrix}$$

$$= \begin{pmatrix} L_2^A \\ L_1^A \\ L_3^A \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 2 & 1 \\ 1 & 2 & 3 \\ 2 & -1 & 0 \end{pmatrix}.$$

1.1.4 Multiplication colonnes par lignes

Cas simple :

Soient A et B deux vecteurs tel que : $A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$ et $B = (b_1, b_2, \cdots, b_n)$.

Alors $AB \in M_{m\times n}(\mathbb{R})$ et on a :

$$AB = \begin{pmatrix} a_1b_1 & \cdots & a_1b_n \\ a_2b_1 & \cdots & a_2b_n \\ \vdots & \vdots & \vdots \\ a_mb_1 & \cdots & a_mb_n. \end{pmatrix}$$

Cas générale :

Soient $A \in M_{m\times n}(\mathbb{R})$ et $B \in M_{n\times p}(\mathbb{R})$, alors $AB \in M_{m\times p}(\mathbb{R})$ et on posant :

$A = (col_1, col_2, \cdots, col_n)$ et $B = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$.

Alors

$$AB = col_1 \times L_1 + col_2 \times L_2 + \cdots + col_n \times L_n.$$

Exemple 2 *Pour* $A, B \in M_2(\mathbb{R})$

$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *et* $B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$, *alors*

$$AB = \begin{pmatrix} a \\ c \end{pmatrix} \begin{pmatrix} e & f \end{pmatrix} + \begin{pmatrix} b \\ d \end{pmatrix} \begin{pmatrix} g & h \end{pmatrix}$$

$$= \begin{pmatrix} ae & af \\ ce & cf \end{pmatrix} + \begin{pmatrix} bg & bh \\ dg & dh \end{pmatrix}$$

$$= \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}.$$

Cas particulier :

Si $A = (a_1, \cdots, a_n)$, alors

$$A^\top = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \quad et \quad AA^\top = \sum_{i=1}^{n} a_i a_i^\top.$$

1.2 Somme directe et espaces supplémentaires

Définition 1 *Soit* E *un* $\mathbb{R}$*-espace vectoriel et* E_1 *et* E_2 *deux sous-espaces vectoriels de* E *on a :*

1. $E_1 + E_2 = \{x + y : x \in E_1 \text{ et } y \in E_2\}$, $E_1 + E_2$ *est un sous-espace vectoriel de* E.
2. *Si* $E_1 \cap E_2 = \{0_E\}$, *on dit que la somme* $E_1 + E_2$ *est directe et on note* $E_1 \oplus E_2$.
3. *Si de plus* $E_1 \oplus E_2 = E$, *on dit que* E_1 *et* E_2 *sont supplémentaires.*

1.2.1 Unicité de la décomposition et propriétés de l'intersection dans le cas de la somme directe

Théorème 1 *Les propositions suivantes sont équivalentes, on a :*

1. $E = E_1 \oplus E_2$.
2. $E_1 \cap E_2 = \{0_E\}$.
3. $\forall x \in E, \exists! a \in E_1, \exists! b \in E_2 : x = a + b$.

Exemple 3 *Dans* $\mathbb{R}^3$, $\{e_1, e_2, e_3\}$ *considérons la base canonique .*

Soit $\Pi = vect\ (e_1, e_2)$ *et* $\triangle = vect\ (e_3)$.

Il est facile de voire que : $\mathbb{R}^3 = \Pi \oplus \triangle$.

Si $\Omega = vect\ \{e_1, e_3\}$ *alors* $\mathbb{R}^3 = \Pi + \Omega$, *mais comme* $\Pi \cap \Omega = vect\ (e_3) \neq \{0_E\}$.

Alors la somme n'est pas directe .

Dans le plan $\mathbb{R}^2$, *si* $\triangle = vect\ ((1,1))$ *et* $D = vect\ ((1,-1))$.

Alors $\mathbb{R}^2 = \triangle \oplus D$.

1.2.2 Bases d'une somme supplémentaire

Théorème 2 *Si* $E = E_1 \oplus E_2$, B_1 *une base de* E_1 *et* B_2 *une base de* E_2 .

Alors

1. $B = B_1 \cup B_2$ *est une base de* E.
2. $dimE = dimE_1 + dimE_2$.

1.2.3 Espaces orthogonaux

Soit u et v deux vecteurs colonnes de $\mathbb{R}^n$. On sait que $u^\top v = (u|v)$ est un produit scalaire . Il est donc possible de définir l'orthogonalité de deux vecteurs :

$$u \perp v = 0 \Leftrightarrow u^\top v = 0.$$

Définition 2 *Soit* U *un sous-espace vectoriel de* $\mathbb{R}^n$.

$$U^\top = \{u \in \mathbb{R}^n : \forall v \in U : u^\top v = 0\}.$$

Autrement dit

$$x \in U^\top \Leftrightarrow (\forall u \in U : (x|u) = 0).$$

Théorème 3

1. $U^\top$ *est un sous-espace vectoriel de* $\mathbb{R}^n$.
2. $\mathbb{R}^n = U \oplus U^\top$.

1.2.4 Projection sur un sous-espace vectoriel dans le cas d'une somme directe

Considérons : $\mathbb{R}^n = U \oplus U^\top$.

On a par définition $\forall x \in \mathbb{R}^n, \exists! a \in U, \exists! b \in U^\top : x = a + b$.

Soit P l'application (linéaire) :

$$\begin{aligned} P : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longrightarrow a \end{aligned}$$

et $P^\perp$ l'application :

$$\begin{aligned} P^\perp : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longrightarrow b \end{aligned}$$

On a

$$\begin{aligned} x \in KerP &\Leftrightarrow P(x) = 0 \\ &\Leftrightarrow a = 0 \\ &\Leftrightarrow x \in U^\top \end{aligned}$$

et

$$\begin{aligned} x \in KerP^\perp &\Leftrightarrow P^\perp(x) = 0 \\ &\Leftrightarrow b = 0 \\ &\Leftrightarrow x \in U. \end{aligned}$$

On a aussi

$$\begin{aligned} y \in ImP &\Leftrightarrow (\exists x \in \mathbb{R}^n : y = P(x)) \\ &\Leftrightarrow y \in U. \end{aligned}$$

Donc $KerP = U^\top$, $ImP = U$, alors $\mathbb{R}^n = KerP \oplus ImP$.

Définition 3 *Si* $\mathbb{R}^n = U \oplus U^\top$ *et* $P : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $x \longrightarrow a$.

Tel que définit plus haut, on dit que P *est la projection sur* U *parallèlement à* $U^\top$.

Théorème 4

1. $\mathbb{R}^n = KerP \oplus ImP = KerP^\perp \oplus ImP^\perp = KerP \oplus KerP^\perp$.
2. $P^2 = P$.
3. $P^\top = P$.
4. $I = P + P^\top$.

Définition 4

On dit que P *est symétrique si* $P^\top = P$.

On dit que P *est idempotent si* $P^2 = P$.

1.3 Détermination des valeurs et vecteurs propres d'une matrice

Définition 5 *On dit que λ est une valeur propre de la matrice A s'il existe un vecteur u non nul solution de*

$$Au = \lambda u.$$

Le vecteur u est le vecteur propre associé à λ.

Théorème 5

λ est une valeur propre de la matrice A si et seulement si

$$det(A - \lambda I) = 0.$$

Le polynôme $det(A - \lambda I)$ est dit polynôme caractéristique de la matrice A, donc trouver les valeurs propres de A, revient à chercher les racines du polynôme caractéristique de la matrice A.

Si $\lambda_1, \lambda_2, \cdots, \lambda_n$ sont les valeurs propres de A, on cherchera les vecteurs propres $u_1, u_2, \cdots, u_n$ respectifs associés aux valeurs propres,qui sont solutions des equations suivantes :

$$Au_i = \lambda u_i \quad ou \quad bien(A - \lambda I)u_i = 0.$$

CHAPITRE

2

ESPACE VECTORIEL ET PROJECTION

Dans ce chapitre, on introduit le concept de matrice de projection. Ce concept joue un rôle important vu qu'il implique le concept "d'orthogonalité" et de "plus courte distance". On parlera de décomposition en somme directe et de supplémentaire orthogonale. Puis en déduira une expression de la matrice projection. On illustrera ce concept par des exemples de projection sur des droits et des plans.
On décrira aussi le principe d'orthogonalisation de Gram-Schmidt qui décrit la manière d'obtenir un système orthonormé à partir d'une famille libre de vecteurs.

2.1 Expression d'une application linéaire

Une application linéaire $f : \mathbb{R}^n \to \mathbb{R}^m$ peut être représentée par une matrice A d'ordre $m \times n$. Une façon de la représenter est de choisir une base orthonormée $\{u_1, \cdots, u_n\}$, puis de calculer les images $a_i = f(u_i)$ et de décomposer ses images dans une base de $\mathbb{R}^m$.

Théorème 6 ***(Fondamental)***[4]

$$A = \sum_{i=1}^{n} a_i u_i^\top.$$

Preuve.

On utilise le fait que A et $\sum_{i=1}^{n} a_i u_i^\top$ sont égaux sur une base, ici on utilise la base orthonormée $\{u_1, \cdots, u_n\}$:

$$Au_i = f(u_i) = a_i = a_i(u_i^\top u_i) = (a_i u_i^\top)u_i.$$

Comme $u_i^\top u_j = \delta_{ij}$ alors

$$Au_i = a_1 u_1^\top u_i + a_2 u_2^\top u_i + \cdots + a_i u_i^\top u_i + \cdots + a_n u_n^\top u_i.$$

Donc

$$Au_i = \sum_{j=1}^{n} (a_j u_j^\top) u_i \quad \forall i = 1, \cdots, n$$

Donc

$$A = \sum_{j=1}^{n} a_j u_j^\top .$$

■

Exemple 4 *Soit* $\{e_1, \cdots, e_n\}$ *la base canonique de* $\mathbb{R}^n$ *on pose* $f(e_j) = a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$, *alors*

$$A = a_1 e_1^\top + \cdots + a_n e_n^\top = \sum_{j=1}^{n} a_j e_j^\top$$

$$A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \begin{pmatrix} 0 & \cdots & 0 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

2.2 Sous-espace et projection

Soit $\{u_1, \cdots, u_r\}$ une famille libre de $\mathbb{R}^n$.

Notons par $U = vect(u_1, \cdots, u_r)$ un sous-espace vectoriel de $\mathbb{R}^n$ de dimension r engendré par cette famille.

Remarques

1- Un sous-espace vectoriel engendré par un seul vecteur non nuls est une droite qui passe par l'origine et qui contient ce vecteur .

2- Un sous-espace vectoriel de $\mathbb{R}^n$ engendré par deux vecteurs non nuls est un plan qui contient l'origine et ces deux vecteurs.

Définition 6 *Soit P un point de $\mathbb{R}^n$ et U un sous-espace vectoriel.Le point Q de U définit de façons que $\overrightarrow{PQ}$ soit orthogonal à U .*

Le point Q s'appelle la projection orthogonale de P sur U (Figure 2.1).

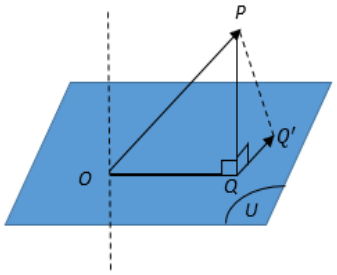


FIGURE 2.1 – La projection orthogonale de P sur U

Théorème 7 *[4]*

La projection orthogonale Q de P sur U vérifie :

$$d(P,U) = \|\overrightarrow{PQ}\|.$$

Autrement dit Q définit la plus petite distance de P au sous-espace U, c'est-à-dire le point de U le plus proche de P.

Preuve.

Soit $Q' \in U$, par la propriété des triangles rectangles qui y apparaissent on a :

$$\|\overrightarrow{PQ'}\| = \|\overrightarrow{PQ} + \overrightarrow{QQ'}\|^2 = \|\overrightarrow{PQ}\|^2 + \|\overrightarrow{QQ'}\|^2 + 2(\overrightarrow{PQ}|\overrightarrow{QQ'}).$$

où $\overrightarrow{PQ} \perp U$ et $\overrightarrow{QQ'} \in U$.

Donc $(\overrightarrow{PQ}|\overrightarrow{QQ'}) = 0$.

Alors

$$\|\overrightarrow{PQ'}\|^2 = \|\overrightarrow{PQ}\|^2 + \|\overrightarrow{QQ'}\|^2 \geq \|\overrightarrow{PQ}\|^2.$$

Finalement

$$\|\overrightarrow{PQ'}\| \geq \|\overrightarrow{PQ}\|.$$

■

2.3 Matrices de projection

$U^\perp$ étant l'ensemble de tous les vecteurs de $\mathbb{R}^n$ orthogonaux a tous les vecteurs de U.

Rappelons aussi que $U^\perp$ est un sous-espace vectoriel de $\mathbb{R}^n$ qui vérifie :

$$\mathbb{R}^n = U \bigoplus U^\perp. \tag{2.1}$$

Il en résulte l'unicité de la décomposition de $\mathbb{R}^n$ en somme direct à partir du sous-espace U.

A partir de (2.1), on peut décomposer n'importe quel vecteur x de $\mathbb{R}^n$:

$$\exists! a \in U; \exists! b \in U^\perp : x = a + b.$$

on a

$$\forall y \in U : \|x - y\|^2 = \|(x - a) + (a - y)\|^2$$

$$= \|b + (a - y)\|^2$$

$$= \|b\|^2 + \|a - y\|^2 + 2(b|a - y).$$

Comme $b \in U^\perp$ et $a - y \in U$, alors $(b|a - y) = 0$.

On déduit :

$$\|x - y\|^2 = \|b\|^2 + \|a - y\|^2$$

$$\geq \|b\|^2 = \|x - a\|^2.$$

donc $\forall y \in U$ on a

$$\|x - y\| \geq \|x - a\|.$$

Comme l'unicité de la projection orthogonale implique l'unicité du point de U qui réalise la distance minimale, on a bien $d(x, U) = a$.

Notons par P l'application :

$$P : \mathbb{R}^n \to U$$
$$x \to a$$

et notons par $P^{\perp}$ l'application :

$$\begin{aligned} P^{\perp} : \mathbb{R}^n &\to U^{\perp} \\ x &\to b \end{aligned}$$

P est l'application linéaire projection orthogonale sur U.

$P^{\perp}$ est l'application linéaire projection orthogonale sur $U^{\perp}$.

On a

$$P(x) = \begin{cases} x & si \quad x \in U \\ 0 & si \quad x \in U^{\perp}. \end{cases}$$

$$P^{\perp}(x) = \begin{cases} 0 & si \quad x \in U \\ x & si \quad x \in U^{\perp}. \end{cases}$$

Car $x = x + 0_{u^{\perp}}$ de façons unique si $x \in U$, et $x = 0_u + x$ de façons unique si $x \in U^{\perp}$.

Soient $\{u_1, \cdots, u_r\}$ une base orthonormée de U.

On peut, en utilisant le théorème de la base incomplète compléter cette base de U en une base orthonormée de $\mathbb{R}^n$ moyennement le procédé de Gram-Schmidt qu'on décrira en fin de chapitre.

Soit $\{u_1, \cdots, u_r, u_{r+1}, \cdots, u_n\}$ cette base orthonormée.

On a

$$P(u_i) = \begin{cases} u_i & si \quad 1 \leq i \leq r \quad car \quad u_i \in U \\ 0 & si \quad r+1 \leq i \leq n \quad car \quad u_i \in U^{\perp}. \end{cases}$$

$$P^{\perp}(u_i) = \begin{cases} 0 & si \quad 1 \leq i \leq r \\ u_i & si \quad r+1 \leq i \leq n. \end{cases}$$

et en utilisant le Théorème 6 on obtient le théorème suivant

Théorème 8 *[4]*

$$\begin{cases} P_r & = u_1 u_1^{\top} + \cdots + u_r u_r^{\top} \\ P_r^{\perp} & = u_{r+1} u_{r+1}^{\top} + \cdots + u_n u_n^{\top}. \end{cases}$$

Corollaire 1

Pour tout sous-espace vectoriel U de $\mathbb{R}^n$, si P_r est la projection orthogonale sur U on a

$$I = P_r + P_r^{\perp}.$$

Preuve.

$$\forall x \in \mathbb{R}^n, \exists! a \in U, \exists! b \in U^{\perp} : x = a + b,$$

où

$$I(x) = x, P_r(x) = a \quad et \quad P_r^{\perp}(x) = b,$$

donc

$$\forall x \in \mathbb{R}^n : I(x) = P_r(x) + P_r^{\perp}(x) = (P_r + P^{\perp})(x),$$

alors

$$I = P_r + P_r^{\perp}.$$

■

Corollaire 2

$$\forall x \in \mathbb{R}^n : \|x\|^2 = \|P_r(x)\|^2 + \|P_r^{\perp}(x)\|^2.$$

Théorème 9

Si P_r est une matrice de projection sur U et $P_r^{\perp}$ la projection orthogonale sur $U^{\perp}$, alors pour tout U on a

$$P_r^{\top} = P_r \tag{2.2}$$

et

$$P_r^2 = P_r. \tag{2.3}$$

Preuve.

1. Pour démontrer la formule (2.2), on obtient

$$P = \sum_{i=1}^{r} u_i u_i^{\top} \Rightarrow P^{\top} = (\sum_{i=1}^{r} u_i u_i^{\top})^{\top}$$

$$= \sum_{i=1}^{r} (u_i u_i^{\top})^{\top}$$

$$= \sum_{i=1}^{r} (u_i^{\top})^{\top} u_i^{\top}$$

$$= \sum_{i=1}^{r} u_i u_i^{\top}.$$

Alors

$$P_r^{\perp} = P_r.$$

On dit que P est symétrique.

2. Pour démontrer la formule (2.3), on a

$$\forall x \in \mathbb{R} : P^2(x) = P(P(x)) \quad ou \quad P(x) \in U,$$

donc $P(P(x)) = P(x)$,

alors

$$P^2 = P.$$

On dit que la projection orthogonale est idempotent $\exists n \in \mathbb{N} : P^n = P$

■

2.4 Projection sur des droites vectorielles et des plans (Hyperplans) vectoriel.

Soit Δ une droite vectorielle, $dim\Delta = 1$ donc il existe un vecteur unitaire

$$u \in \Delta : \Delta = vect(u).$$

Si P est la projection sur Δ alors d'après le théorème 6

$$P = uu^\top.$$

On a

$$\begin{aligned} P(x) &= uu^\top(x) \\ &= u(x|u) \\ &= (x|u)u. \end{aligned}$$

Rappelons que si $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$,

alors

$$\begin{aligned} u^\top v &= \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= \sum_{i=1}^{n} x_i y_i \\ &= (x|y). \end{aligned}$$

Si Π est un hyperplan vectoriel de $\mathbb{R}^n$ ($dim\ \Pi = n-1$), et si n est un vecteur normale ($\| n \| = 1$)

à Π,

alors

$$\begin{cases} \mathbb{R}^n & = \Pi \oplus vect(n) \\ vect(n) & = \Pi^{\perp}. \end{cases}$$

Si P est la projection orthogonal sur Π,

alors

$$I = P + P^{\perp}$$

$$= P + nn^{\top},$$

donc $P = I - nn^{\perp}$ et $P(x) = x - (x|n)n$.

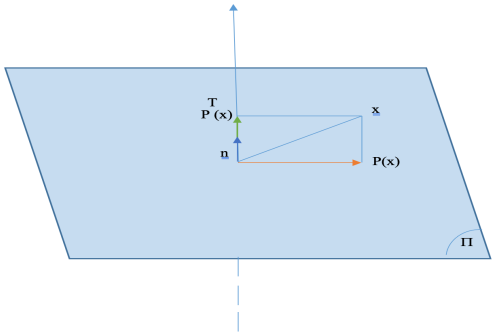


FIGURE 2.2 – La projection orthogonale de P sur un hyperplan vectoriel Π de $\mathbb{R}^n$

2.5 Procédé d'orthogonalisation de Gram-Schmidt

Définition 7

On appelle système orthonormé une famille de vecteurs unitaires deux à deux orthonormaux.

Théorème 10 *[9]*

Il est toujours possible de transformer toute famille libre de vecteurs peut en un système orthonormé.

Preuve.

Soit $\{a_1, \cdots, a_n\}$ une famille libre.

1. Posons $u_1 = \dfrac{a_1}{\|a_1\|}$, donc $\|u_1\| = 1$.

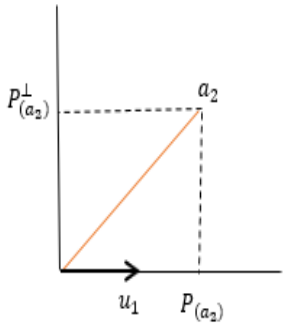


FIGURE 2.3 – Orthonormalisation de Gram-Schmidt

2. On a

$$a_2 = P(a_2) + P^{\perp}(a_2)$$

$$= (a_2|u_1)u_1 + P^{\perp}(a_2),$$

donc

$$P^{\perp}(a_2) = a_2 - (a_2|u_1)u_1.$$

Posons

$$a_2^{'} = P^{\perp}(a_2) = a_2 - (a_2|u_1)u_1. \quad Alors \quad a_2^{'} \perp u_1$$

et posons

$$u_2 = \frac{a_2^{'}}{\|a_2^{'}\|},$$

alors

$$\|u_2\| = 1 \quad et \quad u_1 \perp u_2.$$

3. On a

$$a_3 = P(a_3) + P^{\perp}(a_3).,$$

Où P désigne la projection orthogonale sur $vect(u_1, u_2)$,

alors

$$a_3^{'} = P^{\perp}(a_3)$$

$$= a_3 - (u_1 u_1^{\top} + u_2 u_2^{\top}) a_3$$

$$= a_3 - (a_3|u_1)u_1 - (a_3|u_2)u_2,$$

donc

$$a_3^{'} \perp vect(u_1, u_2)$$

et si $u_3 = \dfrac{a_3^{'}}{\|a_3^{'}\|}$, alors

$$u_3 \perp vect(u_1, u_2), et \|u_3\| = 1.$$

Par itération, on a

$$a_k^{'} = a_k - \sum_{i=1}^{k-1} (a_k|u_i)u_i.$$

$$a_k^{'} \perp vect(u_1, \cdots, u_{k-1}).$$

$$u_k = \frac{a_k^{'}}{\|a_k^{'}\|}.$$

Pour aboutir enfin à un système orthonormal $\{u_1, \cdots, u_n\}$ obtenu à partir de la famille libre $\{a_1, \cdots, a_n\}$. ■

CHAPITRE

3

VALEURS PROPRES ET DÉCOMPOSITION SPECTRALE

On montre dans ce chapitre qu'une matrice symétrique peut s'exprimer en terme de ses valeurs propres et ses vecteurs propres. Cette expression s'appelle la décomposition spectrale. Elle permet de transformer une matrice symétrique en une matrice diagonale par sa multiplication par une matrice orthogonale, à droite et à gauche.
Ce processus s'appelle diagonalisation d'une matrice symétrique. On peut aussi exprimer l'inverse et les puissances d'une matrice symétrique en termes de sa décomposition spectrale.

3.1 Valeurs propres et vecteurs propres

On sait qu'une matrice symétrique d'ordre $n \times n$ est diagonalisable. Donc l'espace $\mathbb{R}^n$ admet une base de vecteurs propres orthonormés. Les valeurs propres sont toutes réelles.
Rappelons que la somme des multiplicités des valeurs propres est égale à $n = dim\mathbb{R}^n$; ce qui signifie que la multiplicité algébrique d'une valeur propre est égale à sa multiplicité géométrique.

Une façon de calculer les valeurs propres serait de calculer les racines du polynôme caractéristique en utilisant une méthode numérique de calcul des racines, comme roots en Matlab ou numpy.roots en Python.
On applique les algorithmes suivants pour trouver les valeurs propres.

Méthode de JACOBI pour matrices réelles symétriques

Méthode de Jacobi pour matrices réelles symétriques Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice réelle symétrique. La méthode de Jacobi (1846) construit une suite de matrices $A = A_0, A_1, \cdots, A_k, \cdots$, qui converge vers une matrice diagonale D.

On a $A_{k+1} = J_k^t A_k J_k$,où J_k est une rotation de Jacobi, $J_k^t J_k = I_n$.

Si l'on pose $J = J_0 J_1 \cdots J_{k-1}$, on aura $D \approx J^t A J$, resp. $A \approx J D J^t$. Comme J est encore une matrice orthogonale, D contient une approximation de toutes les valeurs propres de A et les colonnes de J forment une base orthonormée de vecteurs propres.

Une rotation de Jacobi est définie grâce à une matrice de Givens :

$$G(i,j,\theta) = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & & & & & & \vdots \\ \vdots & & \ddots & & & & & \\ 0 & & & c & & -s & & 0 \\ \vdots & & & & \ddots & & & \vdots \\ 0 & & & s & & c & & 0 \\ \vdots & & & & \ddots & & \ddots & \vdots \\ & & & & & & 1 & 0 \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 1 \end{pmatrix}$$

avec $c = \cos\theta$ et $s = \sin\theta$.

Pour $x \in \mathbb{R}^n$, $G(i,j,\theta)x$ représente une rotation d'angle θ dans le plan $\Pi(e_i, e_j)$.

De même le produit $G(i,j,\theta)^t A$ n'affecte que les i^e etj^e lignes de A, tandis que $AG(i,j,\theta)$ ne transforme que les i^e etj^e colonnes de A.

En particulier, le coût de la multiplication de A par $G(i,j,\theta)$ est en $O(n)$.

Les matrices de Givens sont souvent utilisées pour annuler, par multiplication, un élément donné d'une matrice A.

Dans la méthode de Jacobi, la matrice de Givens est construite grâce à une valeur de θ qui permet d'annuler les éléments a_{ij} et a_{ji}. On montre que θ vérifie alors

$$\cot g(2\theta) = \frac{a_{ii} - a_{jj}}{2a_{ij}} = \tau$$

De cette relation on peut déduire les valeurs de c et s grâce à l'équation vérifiée par

$t = \dfrac{s}{c} = tg\theta, t^2 + 2\tau t - 1 = 0.$

L'algorithme 1 remplace la matrice A par une matrice A' telle que $a'_{ij} = a'_{ji} = 0$.

L'algorithme 1

Rotation de Jacobi pour annuler les éléments de coordonnées $\{i, j\}$

$Si \mid a_{ij} \mid > \epsilon \quad alors$

$$\tau = \frac{a_{ii} - a_{jj}}{2a_{ij}}$$

$$t = \frac{signe(\tau)}{\mid \tau \mid + \sqrt{1 + \tau^2}}$$

$$c = \frac{1}{\sqrt{1 + t^2}} \qquad et \qquad s = ct$$

$$A^{'} = G(i, j, \theta)^t A G(i, j, \theta)$$

Si on note $S(A) = \sum_{1 \leqslant p < q \leqslant n} a_{pq}^2$, alors $S(A^{'}) = S(A) - a_{ij}^2$.

On a ainsi un moyen de mesurer la convergence vers une matrice dont les éléments hors diagonale sont petits. D'où l'algorithme 2 qui converge vers une matrice diagonale contenant les valeurs propres, et une matrice J contenant les vecteurs propres de A.

L'algorithme 2

Algorithme de Jacobi pour valeurs et vecteurs propres d'une matrice symétrique

Seuil $\sigma > 0$ donné, $J = I_n$, $k = 0$

répéter

Choisir $\{i, j\}$ et en déduire $J_k = G(i, j, \theta)$ par l'algorithme 2.13

$A = J_k^t A J^k$
$J = J J_k$
$k = k + 1$
jusqu'à $\quad max_{1 \leq p < q \leq n} \mid a_{pq} \mid < \sigma$

Dans l'algorithme original de Jacobi, on choisit les coordonnées $\{i, j\}$ de façon à avoir $|a_{pq}| = max_{1 \leq p < q \leq n} |a_{kl}|$. Ceci permet d'avoir une bonne vitesse de convergence. Par contre, pour n grand, la recherche du maximum parmi les $(n^2 - n)/2$ éléments hors diagonale va prendre plus de temps par itération que la multiplication par la matrice de Givens. En pratique on parcourt A tout simplement ligne par ligne par exemple.

Exemple

Soit $A = \begin{pmatrix} 2 & 1 & -3 \\ 1 & -6 & 3 \\ -3 & 3 & 0 \end{pmatrix}$ et l'on fixe $\sigma = 10^{-5}$.

On a $spec(A) = \{-7.7155084, -0.5472851, 4.2627936\}$ et $S(A) = 19$.

Pour $J_0 = \begin{pmatrix} 0.8112422 & 0 & 0.847103 \\ 0 & 1 & 0 \\ -0.5847103 & 0 & 0.8112422 \end{pmatrix}$, on annule $a_{13} = a_{31} = -3$:

$A_1 = J_0^t A J_0 = \begin{pmatrix} 4.1622777 & -0.9428887 & 2.2210^{-16} \\ -0.9428887 & -6 & 3.0184368 \\ 2.2210^{-16} & 3.0184368 & -2.1622777 \end{pmatrix}$ avec $S(A_1) = 10$.

Après 5 itérations, on a $J_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.9999904 & -0.0043757 \\ 0 & 0.0043757 & 0.9999904 \end{pmatrix}$ et

$A_5 = J_4^t A_4 J_4 = \begin{pmatrix} 4.2627928 & 0.0029793 & -0.0000130 \\ 0.0029793 & -7.7155077 & 6.9410^{-18} \\ -0.0000130 & 1.2210^{-16} & -0.5472851 \end{pmatrix}$ avec $S(A_5) = 0.0000089$.

En plus $J = J_1 J_2 J_3 J_4 = \begin{pmatrix} 0.7766476 & -0.2219207 & 0.5895504 \\ -0.1060891 & 0.8764440 & 0.4696712 \\ -0.6209376 & -0.4273139 & 0.6571448 \end{pmatrix}$

Méthode de HOUSEHOLDER

Matrice de Householder

Définition 8 *On appelle matrice de Householder, une matrice de la forme :*

$$H = I - 2uu^\top, \quad \text{où} \quad u \in \mathbb{R}^n \quad avec \quad \|u\| = 1.$$

Toute matrice de Householder est symétrique et orthogonale.

La matrice de Householder est la matrice de la symétrie par rapport à l'hyperplan perpendiculaire à u.

L'algorithme de la méthode de Householder

L'algorithme de Householder est une mise en œuvre de la méthode QR différente de l'algorithme de Gram-Schmidt. Il consiste à multiplier la matrice A par une suite de matrices orthogonales très simples (les matrices dites de Householder) pour amener progressivement A sous forme triangulaire supérieure.

A : Données
Début
$H \leftarrow I$
Pour $k = 1, \cdots, n-1$ faire
$\alpha \leftarrow \sqrt{\sum_{i=k}^{m} A_{ik}^2}$
$\beta \leftarrow \alpha^2 - \alpha A_{kk}$
// construction du vecteur v
$v_k \leftarrow A_{kk} - \alpha$
Pour $i = k+1, \cdots, m$ faire
$v_i \leftarrow A_{ik}$
Pour $j = k, \cdots, n$ faire
// Construction de A^{k+1}
$c \leftarrow \frac{1}{\beta} \sum_{i=k}^{m} v_i A_{ij}$
Pour $i = k, \cdots, m$ faire
$A_{ij} \leftarrow A_{ij} - c v_i$
Pour $j = 1, \cdots, m$ faire
// Construction de $H = H_k \cdots H_2 H_1$
$c \leftarrow \frac{1}{\beta} \sum_{i=k}^{m} v_i H_{ij}$
Pour $i = k, \cdots, m$ faire
$H_{ij} \leftarrow H_{ij} - c v_i$
$Q \leftarrow t_H$
Fin
Retourner Q, $R = A$.

3.2 Décomposition Spectrale

Soit $\lambda_1, \cdots, \lambda_n$ les valeurs propres de A et $\{u_i\}$ la base orthonormale des vecteurs propres correspondants, on a

$$Au_i = \lambda_i u_i \quad ; \quad i = 1, \cdots, n$$

Donc A transforme la base a en la base $\{\lambda_i u_i\}$.

Alors d'après le théorème 6, on a

Théorème 11 *[4]*

Pour A est une matrice symétrique et $\{u_i\}$ la base orthonormé formée des vecteurs propres $\{u_i\}$ relatifs aux valeurs propres λ_i, alors

$$A = \lambda_1 u_1 u_1^\top + \cdots + \lambda_n u_n u_n^\top.$$

On sait que $u_i u_i^\top$ est la projection orthogonale sur la direction u_i, donc une matrice symétrique associe a chaque point de $\mathbb{R}^n$ la somme de ses projection orthogonale sur les directions princi-

pales (celles formées par les vecteurs propres) suivie par la multiplication de la valeur propre respective.

Exemple 5 *L'application identité I vérifie :*

$$I = u_1 u_1^\top + \cdots + u_n u_n^\top.$$

Vu que $Iu_i = u_i$, donc toutes les valeurs propres valent 1.

3.3 Décomposition d'une matrice symétrique

Soit $A \in M_n(\mathbb{R})$ telle que $A^\top = A$. Supposons que $rg(A) = r$.

Rappelons que $rg(A)$ est le nombres maximum de vecteurs colonnes de A linéairement indépendant.

Si $a_1, \cdots, a_n$ sont les colonnes de A, alors toute combinaison linéaire des colonnes a_i réalise :

$$\begin{aligned} c_1 a_1 + \cdots + c_n a_n &= \begin{pmatrix} a_1, & \cdots & , a_n \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \\ &= Ac. \end{aligned}$$

Où $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, et $A = \lambda_1 u_1 u_1^\top + \cdots + \lambda_n u_n u_n^\top$.

Donc

$$Ac = \lambda_1 u_1 u_1^\top c + \cdots + \lambda_n u_n u_n^\top c$$

$$Ac = \lambda_1 (u_1|c) u_1 + \cdots + \lambda_n (u_n|c) u_n.$$

Si r est le rang de la matrice A, alors r vecteurs propres sont non nuls, les autres sont nuls ($\lambda_{r+1} = \cdots = \lambda_n = 0$).

Donc

$$Ac = \lambda_1 (u_1|c) u_1 + \cdots + \lambda_r (u_r|c) u_r.$$

Théorème 12

Toute combinaison linéaire des vecteurs colonne de A s'écrit sous la forme d'une combinaison linéaire de r vecteurs mutuellement orthogonaux, donc linéairement indépendants $u_1, \cdots, u_r$.

Corollaire 3

Le rang d'une matrice symétrique est égale au nombres de ses valeurs propres non nulles.

3.4 Diagonalisation des matrices symétriques

L'équation $A = \sum_{i=1}^{n} \lambda_i u_i u_i^\top$ s'écrit :

$$A = \begin{pmatrix} \lambda_1 u_1 & \cdots & \lambda_n u_n \end{pmatrix} \begin{pmatrix} u_1^\top \\ \vdots \\ u_n^\top \end{pmatrix}$$

$$= \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} u_1^\top \\ \vdots \\ u_n^\top \end{pmatrix}$$

$$= U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^\top.$$

Avec $U^\top = \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix}$ est la matrice formée des vecteurs propres de A.

Il est facile de voir que $UU^\top = I$, en effet si on pose $UU^\top = (a_{ij})$ alors

$$a_{ij} = i^{éme} ligne \quad de \quad U \times j^{éme} colonne \quad de \quad U^\top$$

$$a_{ij} = (l_i^\top | c_j^\top).$$

Où l_i est la $i^{éme}$ ligne de U et $c_j^\top$ est $j^{éme}$ colonne de $U^\top$, ainsi la $i^{éme}$ ligne de U est la $i^{éme}$ colonne de $U^\top$.

Donc

$$a_{ij} = (c_i^\top | c_j^\top) = \delta_{ij}.$$

Alors $UU^\top = I$.

Théorème 13 *[9]*

Pour A est une matrice symétrique,

il existe une matrice orthogonale U tel que :

$$U^\top A U = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

U est formé des vecteurs propres de A.

Preuve.[9]

■

3.5 Inverse et Puissance

Soit A une matrice symétrique inversible d'ordre n.

On sait que A est inversible si et seulement si toutes ces valeurs propres non nulles.

Autrement dit : $rang(A) = n$ si $\lambda_1, \cdots, \lambda_n$ les valeurs propres de A.

Théorème 14 *[4]*

$$A^{-1} = \frac{1}{\lambda_1} u_1 u_1^\top + \cdots + \frac{1}{\lambda_n} u_n u_n^\top.$$

Preuve.

Pour λ_i valeur propre de u_i vecteur propre

on a $Au_i = \lambda_i u_i$ et A inversible, donc

$$A^{-1} A u_i = \lambda_i A^{-1} u_i.$$

C'est a dire :

$$A^{-1} u_i = \frac{1}{\lambda_i} u_i.$$

Donc $\frac{1}{\lambda}$ est valeur propre de A^{-1}.

Théorème 15 *[4]*

$$A^{-1} = U \begin{pmatrix} \frac{1}{\lambda_1} & & \\ & \ddots & \\ & & \frac{1}{\lambda_n} \end{pmatrix} U^\top \quad et \quad U^\top A^{-1} U = \begin{pmatrix} \frac{1}{\lambda_1} & & \\ & \ddots & \\ & & \frac{1}{\lambda_n} \end{pmatrix}.$$

Preuve.[4] ■

Théorème 16 *[4]*

Pour tout $p \in \mathbb{N}^*$, on a

$$A^P = \lambda_1^p u_1 u_1^\top + \cdots + \lambda_n^p u_n u_n^\top$$

$$A^P = U \begin{pmatrix} \lambda_1^p & & \\ & \ddots & \\ & & \lambda_n^p \end{pmatrix} U^\top$$

$$U^\top A^P U = \begin{pmatrix} \lambda_1^p & & \\ & \ddots & \\ & & \lambda_n^p \end{pmatrix}.$$

U étant la matrice formé des vecteurs propres de A.

Preuve.[4]

■

Théorème 17 *[4]*

Si f est une fonction polynomiale, alors

$$f(A) = f(\lambda_1)u_1u_1^\top + \cdots + f(\lambda_n)u_nu_n^\top$$

$$f(A) = U \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} U^\top$$

$$U^\top f(A) U = \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix}.$$

U étant la matrice formée de vecteurs propres de A.

Preuve.[4]

■

Définition 9

les valeurs propres de$A \in M_n(\mathbb{R})$ une matrice symétrique sont toutes positives ($\lambda_i \geq 0 \quad i = 1 \cdots n$), alors on peut définir $\sqrt{A}$ comme suit :

$$\sqrt{A} = \sqrt{\lambda_1}u_1u_1^\top + \cdots + \sqrt{\lambda_n}u_nu_n^\top.$$

Théorème 18 *[4]*

$$\sqrt{A} = U \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} U^\top$$

$$U^\top \sqrt{A} U = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}.$$

U étant la matrice formé des vecteurs propres de A.

Remarque 1

Si A est une matrice symétrique définie positive alors A^p peut être définit pour n'importe quel nombre p réel .

Remarque 2

On peut prouver aisément que si on peut définir $\sqrt{A}$ *alors* $\sqrt{A}^2 = A$.

Remarque 3

Pour tout $p \in \mathbb{N}^*$ *:* $(A^{-1})^p = (A^p)^{-1}$.

CHAPITRE

4

VALEURS SINGULIÈRES ET DÉCOMPOSITION EN VALEUR SINGULIER (SVD)

La décomposition spectrale n'est définit que pour les matrices symétriques, donc pour les matrices carrées dans ce chapitre, on étend le principe de la décomposition aux matrice rectangulaires. On définit la décomposition en valeurs singulières, qui permet d'exprimer n'importe quelle matrice en termes de ses valeurs singulières et ses vecteurs singuliers. Les vecteurs singuliers forment une base de l'espace engendré par les vecteurs colonnes, définissant ainsi une matrice de projection sur cet espace.

4.1 Valeurs singuliers et vecteurs singuliers

Soit A une matrice d'ordre $m \times n$ à coefficients réels, alors $A^\top A$ et $AA^\top$ sont deux matrices carrées symétriques d'ordre n et d'ordre m respectivement.
$A^\top A$ et $AA^\top$ symétriques, ces deux matrices considérons une décomposition spectrale :

$$A^\top A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^\top.$$

$$AA^\top = V \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_m \end{pmatrix} V^\top.$$

Où
U étant une matrice orthogonale de vecteurs propres de $A^\top A$.

V une matrice orthogonale de vecteurs propres de $AA^\top$.

Théorème 19 *[8]*

Pour A est une matrice d'ordre $m \times n$ et B une matrice d'ordre $n \times m$, AB et BA ont les mêmes valeurs propres.

Preuve.

Soit $ABx = \lambda x$ et $\lambda \neq 0$ multiplions les deux nombres par B :

$$B(ABx) = \lambda Bx.$$

Étant donné que la multiplication matricielle est associative, alors

$$(BA)(Bx) = \lambda Bx.$$

Ceci veut dire que Bx est vecteur propre de BA pour la même valeur propre λ.

On impose à λ d'être non nul pour être sure que Bx n'est pas nul.

Autrement dit : si $\lambda = 0$ on n'est pas sure que Bx n'est pas nul.

Notons que si B est carrée et inversible, alors

$$B^{-1}(BA)B = AB.$$

Ceci veut dire que BA et AB sont équivalentes donc elles possèdent les même valeurs propre. ■

Corollaire 4 *$A^\top A$ et $AA^\top$ ont même valeurs propres.*

Théorème 20 *Les valeurs propres de $A^\top A$ sont toutes positives.*

Preuve.

Soit λ une valeur propre de $A^\top A$ et u un vecteur propre associée à cette valeur propre

$$A^\top Au = \lambda u.$$

On a

$$(A^\top Au|u) = (Au|Au) = \|Au\|^2.$$

et

$$(A^\top Au|u) = (\lambda u|u) = \lambda\|u\|^2.$$

Donc

$$\|Au\|^2 = \lambda\|u\|^2 \Rightarrow \lambda \geq 0.$$

■

Soit $\lambda_1, \cdots, \lambda_r$ les valeurs propres non nulles de $A^\top A$, on peut les considérer classer par ordre décroissant $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$ et posons $\sigma_i = \sqrt{\lambda_i}$.

Théorème 21 *[10]*

Soit $A \in M_{m\times n}(\mathbb{R})$, alors

$$rg(A^\top A) = rg(A).$$

Preuve.

En montre que $KerA^\top A = KerA$..

Soit $x \in KerA$

donc

$$Ax = 0 \Rightarrow A^\top Ax = 0$$

$$\Rightarrow x \in KerA^\top A.$$

Soit $x \in KerA^\top A$, alors

$$A^\top Ax = 0 \Rightarrow (A^\top Ax|x) = 0$$

$$\Rightarrow \quad (Ax|Ax) = 0$$

$$\Rightarrow \qquad \|Ax\|^2 = 0$$

$$\Rightarrow \qquad\quad Ax = 0$$

$$\Rightarrow x \in KerA,$$

donc

$$KerA^\top A = KerA,$$

alors

$$dimKerA^\top A = dimKerA.$$

Finalement on trouve

$$rg(A^\top A) = rg(A).$$

■

Remarque 4 *:*
Soit $A^\top A \in M_m(\mathbb{R})$.

D'après le théorème du rang on a :

$$m = dimKerA^\top A + rgA^\top A$$

$$\left.\begin{array}{ll} m & = dimKerA + rgA^\top A \\ m & = dimKerA + rgA \end{array}\right\} \Rightarrow rgA^\top A = rgA.$$

Ainsi $\quad rgA = rgA^\top A = rgAA^\top$.

Désignes par r le rang de A.

Uest la matrice formée par les vecteurs propres orthogonaux de $A^\top A$, comme $rgA^\top A = r$, il y r valeurs propres non nulles pour m vecteurs propres distinctes ; $\{u_i\}_i^m$base orthonormée de vecteurs propres de $A^\top A$.

V est la matrice formée par les vecteurs propres orthogonaux de $AA^\top$ et comme $rgAA^\top = r$, il y r valeurs propres non nulles pour n vecteurs propres distinctes ; $\{v_i\}_i^n$base orthonormée de vecteurs propres de $AA^\top$.

Les valeurs propres de $AA^\top$ non nulles sont $\sigma_1^2, \cdots, \sigma_r^2$, les autres sont nulles.

et

$$A^\top A = \sum_{i=1}^{r} \sigma_i^2 a_i u_i^\top \qquad a_i = A^\top A u_i.$$

$$AA^\top = \sum_{i=1}^{r} \sigma_i b_i v_i^\top.$$

Posons $\sigma_i^2 = \lambda_i$ et $u_i = \dfrac{1}{\sigma_i} A v_i \in \mathbb{R}^m$, tels que $\sigma_i \neq 0 \quad \forall 1 \leq i \leq r$.

Théorème 22 *[8]*

$$(u_i \mid u_j) = \delta_{ij} \qquad \forall 1 \leq i, j \leq n.$$

Preuve.

$$(u_i \mid u_j) = (\frac{1}{\sigma_i} Av_i \mid \frac{1}{\sigma_j} Av_j)$$

$$= \frac{1}{\sigma_i \sigma_j} (Av_i \mid Av_j)$$

$$= \frac{1}{\sigma_i \sigma_j} (\sigma_i u_i \mid \sigma_j u_j)$$

$$= \frac{\sigma_i \sigma_j}{\sigma_i \sigma_j} (u_i \mid u_j).$$

Alors

$$(u_i \mid u_j) = \delta_{ij}.$$

■

Remarque 5

$$r = rang(A) \leq \min(m, n).$$

Il y a r vecteurs u_i *définit par* $u_i = \frac{1}{\sigma_i} Av_i$.

On peut choisir $r - m$ u_i *autres vecteurs de* $\mathbb{R}^m$ *deux a deux orthogonaux de façons que* $\{u_i\}_1^m$ *soient une bases orthonormées (Gram-Schmidt).*

Théorème 23 *[8]*

$$Av_i = \sigma_i u_i \qquad (par \quad définition)$$

$$A^\top u_i = \sigma_i v_i \qquad (1 \leq i \leq n)$$

Preuve.

$$A^\top u_i = \frac{1}{\sigma_i} A^\top (Av_i)$$

$$= \frac{1}{\sigma_i} (A^\top A) v_i$$

$$= \frac{1}{\sigma_i} \sigma_i^2 v_i,$$

alors on trouve que

$$A^\top u_i = \sigma_i v_i \qquad (1 \leq i \leq n)$$

■

Théorème 24

$$A = \sigma_1 u_1 v_1^\top + \cdots + \sigma_r u_r v_r^\top.$$

$$A^\top = \sigma_1 v_1 u_1^\top + \cdots + \sigma_r v_r u_r^\top,$$

telles que $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$

Preuve. Application directe du Théorème 6.

Définition 10

Les σ_i s'appellent les valeurs singuliers de la matrice A.

Les v_i s'appellent les vecteurs singuliers droit de la matrice A.

Les u_i s'appellent les vecteurs singuliers gauche de la matrice A.

4.2 Espace colonne et espace ligne d'une matrice

Soit $A \in M_{m\times n}(\mathbb{R})$, possède n colonnes et m lignes.

On a

$$U = vect(col_1, col_2, \cdots, col_n) = Im(A) \subset \mathbb{R}^m.$$

$$V = vect(lig_1, lig_2, \cdots, lig_n) = Im(A^\top) \subset \mathbb{R}^n.$$

Alors

$$dimU = dimV = rang \quad de \quad A = r.$$

Posons

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^\top \quad o\grave{u} \quad c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n.$$

Donc

$$Ac = \sum_{i=1}^{r} \sigma_i u_i v_i^\top c$$

$$= \sum_{i=1}^{r} \sigma_i u_i (c|v_i)$$

$$= \sigma_1 (c|v_1) u_1 + \cdots + \sigma_r (c|v_r) u_r.$$

Donc

$U = vect(u_1, \cdots, u_r)$ et comme $u_1, \cdots, u_r$ sont deux à deux orthogonaux, alors

$\{u_1, \cdots, u_r\}$ est une base de U.

Il en est de même pour V

$$V = vect(v_1, \cdots, v_r).$$

Théorème 25 *[4]*

$\{v_i\}_i^r$ est une base orthonormée de $Im(A^\top) = V$.

$\{u_i\}_i^r$ est une base orthonormée de $Im(A) = U$.

Théorème 26 *[4]*

$$P_U = u_1 u_1^\top + \cdots + u_r u_r^\top.$$

$$P_V = v_1 v_1^\top + \cdots + v_r v_r^\top.$$

Preuve.

Application directe du Théorème 6 et la définition de la projection sur un espace.

4.3 Représentation matricielle

On a

$$A = \sigma_1 u_1 v_1^\top + \cdots + \sigma_r u_r v_r^\top$$

$$= \begin{pmatrix} \sigma_1 u_1 & \cdots & \sigma_r u_r \end{pmatrix} \begin{pmatrix} v_1^\top \\ \vdots \\ v_r^\top \end{pmatrix}$$

$$= \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \begin{pmatrix} v_1^\top \\ \vdots \\ v_r^\top \end{pmatrix}.$$

et comme on a $U = \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix}$ et $V^\top = \begin{pmatrix} v_1^\top \\ \vdots \\ v_r^\top \end{pmatrix}$.

Alors

$$A = U \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} V^\top.$$

où $U \in M_{m\times r}(\mathbb{R})$ et $V \in M_{r\times n}(\mathbb{R})$.

Théorème 27 *[4]*

(1) $U^\top U = I$.

(2) $V^\top V = I$.

(3) $UU^\top = P_U$.

(4) $VV^\top = P_V$.

Preuve.

Pour démontrer la formule (1), on prend $U = \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix}$ et $U^\top = \begin{pmatrix} u_1^\top \\ \vdots \\ u_r^\top \end{pmatrix}$.

On a $(u_i|u_j) = \delta_{ij}$.

En effet, si on pose $U^\top U = (\alpha_{ij})$.

Alors

$$\begin{aligned} \alpha_{ij} &= i^{iéme} ligne \quad de \quad U^\top \times j^{iéme} colonne \quad de \quad U \\ &= (i^{iéme} colonne \quad de \quad U | j^{iéme} colonne \quad de \quad U) \\ &= (u_i|u_j). \end{aligned}$$

Donc

$$\alpha_{ij} = \delta_{ij}.$$

De même manière on montre la formule (2).

Pour démontrer la formule (3), on a $U = \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix}$ et $U^\top = \begin{pmatrix} u_1^\top \\ \vdots \\ u_r^\top \end{pmatrix}$.

Alors

$$\begin{aligned} UU^\top &= \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix} \begin{pmatrix} u_1^\top \\ \vdots \\ u_r^\top \end{pmatrix} \\ &= \sum_{i=1}^{r} u_i u_i^\top. \end{aligned}$$

Donc on trouve

$$UU^\top = P_U.$$

De la même manière on montre la formule (4).

4.4 Pseudo-inverse

Définition 11

Soit A une matrice telles que $A \in M_{m\times n}(\mathbb{R})$, possède une décomposition en valeurs singuliers.

Où

$$A = \sigma_1 u_1 v_1^\top + \cdots + \sigma_r u_r v_r^\top .pour \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

On appelle "Pseudo-inverse" on inverse généralisé, la matrice suivante :

$$A^- = \frac{v_1 u_1^\top}{\sigma_1} + \cdots + \frac{v_r u_r^\top}{\sigma_r}.$$

Théorème 28 *[4]*

$$A^- = V \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_r} \end{pmatrix} U^\top.$$

D'où

$$V = \begin{pmatrix} v_1 & \cdots & v_r \end{pmatrix} \quad et \quad U = \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix}.$$

v_i vecteur propre orthonormé de $A^\top A$.

$u_i = \dfrac{Av_i}{\sigma_i} \quad 1 \leq i \leq r.$

Théorème 29 *[4]*

$$A^- A = P_V = la \quad projection \quad sur \quad col(A^\top). \tag{4.1}$$

$$AA^- = P_U = la \quad projection \quad sur \quad col(A). \tag{4.2}$$

Preuve.

Pour démontrer la formule (4.1), on a

$$A^- A = (\sum_{i=1}^{r} \frac{v_i u_i^\top}{\sigma_i})(\sum_{i=1}^{r} \sigma_i u_i v_i^\top)$$

$$= \sum_{1 \leq i,j \leq r} \frac{v_i u_i^\top}{\sigma_i} \sigma_j u_j v_j^\top$$

$$= \sum_{1 \leq i,j \leq r} \frac{\sigma_j}{\sigma_i} v_i u_i^\top u_j v_j^\top$$

$$= \sum_{1 \leq i,j \leq r} \frac{\sigma_j}{\sigma_i} v_i (u_i | u_j) v_j^\top$$

$$= \sum_{1 \leq i,j \leq r} \delta_{ij} \frac{\sigma_j}{\sigma_i} v_i v_j^\top$$

$$= v_1 v_1^\top + \cdots + v_r v_r^\top.$$

Donc on trouve que

$$A^- A = P_V.$$

De même manière on montre la formule(4.2).

Remarque 6

$$\forall x \in U : \quad P_U x = x.$$

Donc P_U joue le rôle de la matrice identité dans l'espace col(A).

$$\forall x \in V : \quad P_V x = x.$$

Donc P_V joue le rôle de la matrice identité dans l'espace col$(A^\top)$ = espace ligne de A.

D'où la définition du nom pseudo-inverse.

Théorème 30 *[4]*

$$P_V A^- = A^-. \tag{4.3}$$

$$A^- P_U = A^-. \tag{4.4}$$

D'où $P_V = v_1 v_1^\top + \cdots + v_r v_r^\top$ et $P_U = u_1 u_1^\top + \cdots + u_r u_r^\top$.

Preuve.

1. Pour démontrer la formule (4.3), on a

$$A^- = \frac{v_1 u_1^\top}{\sigma_1} + \cdots + \frac{v_r u_r^\top}{\sigma_r}.$$

$$P_V A^- = \frac{P_V v_1 u_1^\top}{\sigma_1} + \cdots + \frac{P_V v_r u_r^\top}{\sigma_r} \qquad telles \quad que(v_i \in V).$$

$$= \frac{v_1 u_1^\top}{\sigma_1} + \cdots + \frac{v_r u_r^\top}{\sigma_r}.$$

   Alors on trouve que

$$P_V A^- = A^-.$$

2. Pour démontrer la formule (4.4), on a d'après le Théorème 29

$$A A^- = P_U.$$

   Donc

$$A^- A A^- = A^- P_U.$$

Où

$$(A^-A)A^- = P_V A^- = A^-.$$

Finalement on trouve

$$A^- P_U = A^-.$$

■

Théorème 31 *[4]*

$$AA^-A = A. \tag{4.5}$$

$$A^-AA^- = A^-. \tag{4.6}$$

Preuve.

1. Pour démonter la formule (4.5), on a

$$AA^-A = (AA^-)A.$$

D'après le Théorème 29

$$AA^-A = P_U A.$$

Alors

$$AA^-A = P_U(\sum_{i=1}^{r} \sigma_i u_i v_i^\top).$$

Donc

$$AA^-A = \sum_{i=1}^{r} \sigma_i P_U u_i v_i^\top = \sum_{i=1}^{r} \sigma_i u_i v_i^\top \qquad car \quad (u_i \in U).$$

Alors on trouve

$$AA^-A = A.$$

2. Pour prouver la formule (4.6), on utilise le Théorème 29

$$A^-A = P_V.$$

Alors

$$P_V A^- = A^-.$$

Donc

$$A^-AA^- = A^-.$$

■

4.5 Pseudo inverse d'un vecteurs

Soit $a \in \mathbb{R}^n$; $a \in M_{n\times 1}(\mathbb{R})$, donc a admet un pseudo -inverse.

On a

$$a^- = \frac{1}{\|a\|} \cdot 1 \cdot \frac{a^\top}{\|a\|}$$

$$= \frac{a^\top}{\|a\|^2}.$$

et

$$a = \sum_{i=1}^{n} \sigma_i u_i v_i^\top$$

$$= \|a\| \cdot (\frac{a}{\|a\|}) \cdot 1.$$

Alors

$$a^- a = \frac{a^\top}{\|a\|^2} \cdot \|a\| \cdot \frac{a}{\|a\|} \cdot 1 = 1.$$

$$aa^- = \|a\| \cdot (\frac{a}{\|a\|}) \cdot 1 \cdot \frac{a^\top}{\|a\|^2} = \frac{aa^\top}{\|a\|^2}.$$

Posons $u = \frac{a}{\|a\|}$, alors on a

$$aa^- = a \cdot \frac{a^\top}{\|a\|^2}$$

$$= \frac{a}{\|a\|} \cdot \frac{a^\top}{\|a\|}.$$

Alors

$$aa^- = uu^\top.$$

et

$$a^- a = \frac{a^\top}{\|a\|^2} \cdot a$$

$$= \frac{(x|a)}{\|a\|^2}.$$

Donc

$$a^- a = 1.$$

$a^- a$ est la projection sur $\mathbb{R}$.

aa^- est la projection sur u.

CHAPITRE

5

ESPACES D'AJUSTEMENTS

5.1 Droite d'ajustement

Considérons le problème qui consiste à trouver une droite vectorielle approximante.

Soit v une base unitaire ($\|v\| = 1$) de cette droite.

La longueur de la projections de x_α sur la direction de v est donnée par $(x_\alpha \mid v)$.

Ainsi la somme des carrés de ces longueurs est donnée par

$$k = \sum_{\alpha=1}^{N} (x_\alpha \mid v)^2$$

$$= \sum_{\alpha=1}^{N} v^\top x_\alpha x_\alpha^\top v$$

$$= (v \mid \sum_{\alpha=1}^{N} x x_\alpha^\top v)$$

$$= (v \mid (\sum_{\alpha=1}^{N} x_\alpha x_\alpha^\top) v),$$

alors

$$K = (v \mid \Sigma v),$$

où $\Sigma = \sum_{\alpha=1}^{N} x_\alpha x_\alpha^\top$.

Σ est une matrice $n \times n$. Cette matrice s'appelle matrice des moments dite aussi matrice de covariance.

5.2 Espace d'ajustement

Soient N points $x_1, \cdots, x_N \in \mathbb{R}^n$.

On considère le problème qui consiste à trouver un sous-espace dans $\mathbb{R}^r$ tel qu'il soit le plus proche possible de ces points, on suppose que $N \geq r$.

Si $n = 3$ et $r = 1$, le problème est de trouver une droite passant par l'origine et qui soit le plus proche possible de ces points.

Si $n = 3$ et $r = 2$, le problème est de trouver un plan vectoriel qui soit le plus proche possible de ces points.

Ici le notion de proximité (proche) est définit par la somme des carrés des distances.

Trouver un sous-espace c'est trouver une base à ce sous-espace.

Soit $(u_1, \cdots, u_r)$ une base orthonormée de sous-espace approximant U, et $(u_1, \cdots, u_n)$ une base orthonormée de $\mathbb{R}^n$.

Autrement dit, le problème est de trouver une base orthonormée de $\mathbb{R}^n$ de façons que le sous-espace U engendré par $(u_1, \cdots, u_r)$ est le meilleur espace approximant au sens de la distance euclidienne.

On sait que

$$d(x_\alpha, U) = \| P^\perp x_\alpha \| .$$

et que

$$P^\perp = u_{r+1}u_{r+1}^\top + \cdots + u_n u_n^\top.$$

Donc la somme des distances des N points $x_1, \cdots, x_N$ au sous-espace U est donnée par J.

Telle que

$$J = \sum_{\alpha=1}^{N} \|P^\perp x_\alpha\|^2.$$

On veut donc minimiser J, on a

$$\|x_\alpha\|^2 = \|Px_\alpha\|^2 + \|P^\perp x_\alpha\|^2.$$

D'où

$$\|P^\perp x_\alpha\|^2 = \|x_\alpha\|^2 - \|Px_\alpha\|^2.$$

Ainsi

$$J = \sum_{\alpha=1}^{N} (\|x_\alpha\|^2 - \|Px_\alpha\|^2).$$

5.2.1 Extremums d'une forme quadratique

Théorème 32

Soit $A \in M_n(\mathbb{R})$ est une matrice symétrique.

Soit la forme quadratique (v, Av) où v est un vecteur unitaire ($\|v\| = 1$).

Soient $\lambda_1, \cdots, \lambda_n$ les valeurs propres de A tel que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$,et soit $\{u_i\}_{i=1}^n$ la base orthonormée des vecteurs propres associés à ces valeur propre.

Quels que soit $v \in \mathbb{R}^n$:

$$v = \sum_{i=1}^{n} c_i u_i \quad et \quad \sum_{i=1}^{n} c_i^2 = 1.$$

Preuve.

On sait que

$$(v \mid v) = (\sum_{i=1}^{n} c_i u_i \mid \sum_{i=1}^{n} c_i u_i)$$

$$= \sum_{1 \leq i,j \leq n} c_i c_j (u_i \mid u_j).$$

Alors

$$(v \mid v) = \sum_{i=1}^{n} c_i^2.$$

■

Théorème 33 *[4]*

Soit $A \in M_n(\mathbb{R})$ et $v \in \mathbb{R}_n$:

$$(v \mid Av) = \sum_{i=1}^{n} \lambda_i c_i^2.$$

Preuve.

On sait que

$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^\top.$$

Donc

$$Av = \sum_{i=1}^{n} \lambda_i (v \mid u_i) u_i$$

$$= \sum_{i=1}^{n} \lambda_i (\sum_{j=1}^{n} c_j u_j \mid u_i) u_i.$$

On trouve que

$$Av = \sum_{i=1}^{n} \lambda_i c_i u_i.$$

Alors

$$(v \mid Av) = (\sum_{i=1}^{n} c_i u_i \mid \sum_{i=1}^{n} \lambda_i c_i u_i)$$

$$= \sum_{1 \leq i,j \leq n} c_i \lambda_j c_j (u_i \mid u_j).$$

Finalement on a

$$(v \mid Av) = \sum_{i=1}^{n} \lambda_i c_i^2.$$

■

Corollaire 5 *[4]*

$$Max(v \mid Av) = \lambda_1.$$

$$Max(v \mid Av) = (u_1 \mid Au_1).$$

$$Min(v \mid Av) = \lambda_n.$$

$$Min(v \mid Av) = (u_n \mid Au_n).$$

Preuve.

D'après le Théorème 33, on a

$$(v \mid Av) = \sum_{i=1}^{n} \lambda_i c_i^2 \leq \sum_{i=1}^{n} \lambda_1 c_i^2.$$

Alors

$$(v \mid Av) = \lambda_1 \sum_{i=1}^{n} c_i^2 \leq \lambda_1.$$

et on a

$$(u_1 \mid Au_1) = \sum_{i=1}^{n} \lambda_i c_i^2 \quad o\grave{u} \quad u_1 = 1 \cdot u_1 + 0 \cdot u_2 + \cdots + 0 \cdot u_n.$$

Donc

$$(u_1 \mid Au_1) = \lambda_1 \quad et \quad Max(v \mid Av) = \lambda_1 = (u_1 \mid Au_1).$$

■

5.2.2 Espace d'ajustements hiérarchisés

Lemme 1 *:[11]*

Soit $u \in M_{n\times 1}(\mathbb{R})$ *alors*

$$Tr(uu^\top) = \|u\|^2.$$

Preuve.

$$uu^\top = (\alpha_{ij})$$

$$\alpha_{ii} = u_i u_i = u_i^2.$$

Donc

$$Tr(uu^\top) = \sum_{i=1}^{n} \alpha_{ii}$$

$$= \sum_{i=1}^{n} u_i^2$$

$$= \|u\|^2.$$

Alors

$$Tr(uu^\top) = \|u\|^2.$$

■

Corollaire 6

$$Tr(\Sigma) = \sum_{\alpha=1}^{N} \|x_\alpha\|^2.$$

Preuve.

$$Tr(\Sigma) = Tr(\sum_{\alpha=1}^{N} x_\alpha x_\alpha^\top)$$

$$= \sum_{\alpha=1}^{N} Tr(x_\alpha x_\alpha^\top).$$

Alors

$$Tr(\Sigma) = \sum_{\alpha=1}^{N} \|x_\alpha\|^2.$$

■

La direction u_1 ainsi obtenue indique l'orientation dans la quelle les N points $\{x_\alpha\}$ se repartissent dans la plus grande distance.

Si toutefois les points se repartissent dans une autre direction alors cette approximation n'est pas suffisante.
On cherche donc une autre direction v qui soit orthogonale à u_1 le long de la quelle la somme des carrés des projections soit maximale.

Cette somme est donnée par K tel que :

$$K = (v|\Sigma v).$$

et on a résoudre le problème :

$$\begin{cases} \max K \\ \\ (v \mid u_1) = 0. \end{cases}$$

Théorème 34 *[4]*
u_2 est la solution au problème précédent.

Preuve.

$$K = \sum_{i=1}^{n} \sigma_i^2 c_i^2$$

$$K \leq \sigma_1^2 c_1^2 + \sigma_2^2 \sum_{i=2}^{n} c_i^2.$$

Si $v = \sum\limits_{i=1}^{n} c_i u_i$ et $(v \mid u_1)$, alors $c_1 = 0$.

et comme $\|v\| = 1$ donc $\sum\limits_{i=2}^{n} c_i^2 = 1$.

Alors

$$K \leq \sigma_2^2.$$

Si on choisit $v = u_2$ alors $K = \sigma_2^2$.

Donc

$$\max K = \sigma_2^2$$

$$\max K = (u_2 \mid \Sigma u_2).$$

■

Théorème 35

Si $U = vect(u_1, u_2)$ *et* $P = P_U$ *alors*

$$K = \sigma_1^2 + \sigma_2^2$$

.

Preuve.

$$K = \sum_{\alpha=1}^{N} \|P_{x_\alpha}\|^2.$$

$$P = u_1 u_1^\top + u_2 u_2^\top.$$

$$P_{x_\alpha} = (x_\alpha \mid u_1)u_1 + (x_\alpha \mid u_2)u_2.$$

$$\|P_{x_\alpha}\|^2 = (x_\alpha \mid u_1)^2 + (x_\alpha \mid u_2)^2.$$

Donc

$$K = \sum_{\alpha=1}^{N} (x_\alpha \mid u_1)^2 + \sum_{\alpha=1}^{N} (x_\alpha \mid u_2)^2 < \sigma_1^2 + \sigma_2^2.$$

Donc

$$\max K = \sigma_1^2 + \sigma_2^2.$$

■

Théorème 36 *[4]*

Si $U = vect(u_1, \cdots, u_r)$ *et* $P = P_U$ *on a*

$$J = min(\sum_{\alpha=1}^{N} \|x_\alpha\|^2 - K).$$

Alors

$$J = \sigma_{r+1}^2 + \cdots + \sigma_n^2.$$

Preuve.

$$\sum_{\alpha=1}^{N} \|x_\alpha\|^2 = Tr(\Sigma)$$

$$= \sigma_1^2 + \cdots + \sigma_n^2.$$

Alors

$$K = \sigma_1^2 + \cdots + \sigma_r^2.$$

Donc

$$J = \sum_{i=1}^{N} \sigma_i^2 - \sum_{i=1}^{r} \sigma_i^2$$

On trouve que

$$J = \sum_{r+1=1}^{n} \sigma_i.$$

■

J s'appelle la somme résiduelle.

CONCLUSION

Dans le vaste domaine de la factorisation matricielle, nous avons présenté dans ce mémoire le concept de matrice de projection, qu'il implique le concept "d'orthogonalité" et de "plus courte distance ».

Le calcul des valeurs propres et vecteurs propres de matrices jouent des rôles importants dans de nombreuses applications. Nous avons déduit le théorème de la décomposition spectrale, qu'elle permet de transformer une matrice symétrique en une matrice diagonale.

En effet, nous avons également vu que cette décomposition spectrale n'est définie que pour les matrices symétriques donc pour les matrices carrées on définit la décomposition en valeurs singulières. Ainsi, nous avons déduit l'étude des espaces d'ajustements aux dimension supérieure.

BIBLIOGRAPHIE

[1] G. Allaire, et S.M. Kaber,(2002). Algebre linéaire numérique.Ellipses.

[2] J.Bisgard, (2021).Analysis and Linear Algebra : the singular value decomposition and Applications ;Studenet Mathematical Library, vol(94), Aerican Mathematical Society, Providence, Rhode Island.

[3] J. Grifone, (2019). Algèbre Linéaire 6E Édition. Éditions Cépaduès.

[4] K.Kanatani. (2021). Linear Algebra for Pattern processing, projecton, singular value decomposition and pseudoinverse. José Moura, Carnegie Mellon University

[5] C.Lay and R.Lay and J.Mcdonald, (2015). LInear Algebra and its applications. Pearson

[6] C. Lay, (2004). Algèbre linéaire : Théorie, exercices et applications. De Boeck Supérieur.

[7] E.Matzner-Lober,(2007). Regression, théorie et application (statistique et probabilité appliquée) ; Springer Verlag, France, Paris.

[8] S.Gilbert,(2019) Linear Algebra and learning from data, Wellesley Cambridge Press, Wellesley MA.

[9] S.Gilbert,(2016). Introduction to linear Algebra. Wellesley Cambridge Press, Wellesley MA.

[10] https ://math.stackexchange.com/questions/349738/proveoperatornamerankata-operatornameranka-for-any-a-in-m-m-times-n

[11] https ://math.stackexchange.com/questions/2452972/what-can-trace-of-tt-say-about-t.