ALGÈBRE POUR LA STATISTIQUE
Diaporama du Polycopié

FSR - LMSA
Master : Statistiques et Econométrie
Nabil Machrafi
Années universitaires : 2018-2019 et 2019-2020

Les chapitres

- Chapitre préliminaire
- Espaces euclidiens et orthogonalité
- Décompositions spectrale et en valeurs singulières

Référence : Linear Algebra and Matrix Analysis for Statistics, Sudipto Banerjee et Anindya Roy, 2014.

Produit matriciel

Définition 1

Soient $A = (a_{ij})$ *et* $B = (b_{ij})$ *deux matrices de taille* (m, n) *et* (n, p) *respectivement. Le produit des matrices A et B est la matrice de taille* (m, p) *définie par :*

$$AB = (c_{ij}) \text{ avec } c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \text{ , } 1 \leq i \leq m, \ 1 \leq j \leq p.$$

Cas particuliers

Produit tensoriel de deux vecteurs colonnes $x = {}^t(x_1, ..., x_m)$ et $y = {}^t(y_1, ..., y_n)$:

$$\begin{aligned} x \otimes y \quad : & = x\, {}^t y \\ & = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \\ & = \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & & \vdots \\ x_m y_1 & \cdots & x_m y_n \end{bmatrix} = \left(x_i y_j\right)_{\substack{1 \le i \le m \\ 1 \le j \le n}} . \end{aligned}$$

Cas particuliers

Produit scalaire (canonique) sur $\mathbb{R}^n$:
Pour $x = {}^t(x_1, ..., x_n)$ et $y = {}^t(y_1, ..., y_n)$, le produit scalaire de x et y est le scalaire :

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

Différentes expressions du produit matriciel

via le produit scalaire :

$$AB = \begin{bmatrix} {}^t a_{1*} \\ \vdots \\ {}^t a_{m*} \end{bmatrix} \begin{bmatrix} b_{*1} & \cdots & b_{*n} \end{bmatrix} = \begin{bmatrix} {}^t a_{1*} b_{*1} & \cdots & {}^t a_{1*} b_{*n} \\ \vdots & & \vdots \\ {}^t a_{m*} b_{*1} & \cdots & {}^t a_{m*} b_{*n} \end{bmatrix}.$$

via le produit tensoriel :

$$AB = \begin{bmatrix} a_{*1} & \cdots & a_{*p} \end{bmatrix} \begin{bmatrix} {}^t b_{1*} \\ \vdots \\ {}^t b_{p*} \end{bmatrix} = \sum_{i=1}^{p} a_{*i} \otimes b_{i*}.$$

Différentes expressions du produit matriciel

Par concaténation de vecteurs colonnes :

$$AB = A\begin{bmatrix} b_{*1} & \cdots & b_{*n} \end{bmatrix} = \begin{bmatrix} Ab_{*1} & \cdots & Ab_{*n} \end{bmatrix}.$$

Par concaténation de vecteurs lignes :

$$AB = \begin{bmatrix} {}^t a_{1*} \\ \vdots \\ {}^t a_{m*} \end{bmatrix} B = \begin{bmatrix} {}^t a_{1*} B \\ \vdots \\ {}^t a_{m*} B \end{bmatrix},$$

Théorème du rang

Théorème 1

Soit A une matrice de taille (m, n) . *Alors*

$$rg\,(A) + \dim KerA = n,$$

où KerA est le sous espace vectoriel noyau de A :

$$KerA := \{x \in \mathbb{R}^n : Ax = 0\}\,.$$

Corollaire 1

Une matrice carrée A de taille n est inversible si et seulement si $rg\,(A) = n$.

Théorème du rang

Exercice 1

Soit A une matrice de taille (m, n) .

1. *Montrer que $Ker({}^tAA) = KerA$.*
2. *En déduire que*

$$rg\,(A) = rg({}^tAA) = rg(A\,{}^tA).$$

Matrices de rang plein

Définition 2

Une matrice A de taille (m, n) est dite de rang plein si $rg\,(A) = n$.

D'après le théorème du rang, une matrice A est de rang plein si et seulement si $\mathrm{Ker}A = \{0\}$, i.e.

$$Ax = 0 \Rightarrow x = 0.$$

Proposition 1

Soit A une matrice de taille (m, n) . Alors

$$A \text{ est de rang plein } \Leftrightarrow {}^{t}AA \text{ est inversible.}$$

Trace d'une Matrice

Définition 3

Soit $A = (a_{ij})$ une matrice carrée de taille n. La trace de A, notée trA, est la somme de ses coefficients diagonaux :

$$trA = \sum_{i=1}^{n} a_{ii}.$$

Proposition 2 (Propriété cyclique de la trace)

Soient A et B deux matrices de tailles (m, n) et (n, m) respectivement. Alors

$$tr(AB) = tr(BA).$$

Trace d'une Matrice

Exercice 2

1. *Montrer que pour* $x, y \in \mathbb{R}^n$, $tr(x\,{}^t y) = {}^t x\, y$.
2. *Soient* A *et* B *deux matrices de tailles* (m, n) *et* (n, m) *respectivement. Montrer que* $tr\,(AB) = tr\,(BA)$.

Projecteurs

Définition 4

1. *Soit $\mathbb{R}^n = E \oplus F$. Une matrice carrée P de taille n est dite matrice de projection sur E parallèlement à F, si pour tout $x \in \mathbb{R}^n$, Px est la projection de x sur E parallèlement à F.*
2. *Une matrice carrée P de taille n est dite un projecteur s'elle est une matrice de projection sur un sous espace vectoriel de $\mathbb{R}^n$ parallèlement à un supplémentaire.*

Projecteurs

Théorème 2

Pour une matrice carrée P de taille n, les assertions suivantes sont équivalentes :

1. *P est un projecteur ;*
2. *P est une matrice idempotente, i.e. $P^2 = P$;*
3. *$\mathbb{R}^n = ImP \oplus KerP$ et P est la matrice de projection sur ImP parallèlement à $KerP$.*

Matrices par blocs

Proposition 3

Soit A une matrice de taille (m, n) partitionnée en une matrice par blocs de la forme

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \tag{1}$$

Alors la transposée de A est donné par

$$^tA = \begin{bmatrix} ^tA_{11} & ^tA_{21} \\ ^tA_{12} & ^tA_{22} \end{bmatrix}$$

Matrices par blocs

On dit que deux matrices par blocs $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ et $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ sont conformément partitionnées pour l'addition si leurs blocs correspondants sont des matrices de même taille.
Dans ce cas, pour $\alpha, \beta \in \mathbb{R}$ on a

$$\alpha A + \beta B = \begin{bmatrix} \alpha A_{11} + \beta B_{11} & \alpha A_{12} + \beta B_{12} \\ \alpha A_{21} + \beta B_{21} & \alpha A_{22} + \beta B_{22} \end{bmatrix}.$$

Déterminant d'une matrice par blocs

Théorème 3

Soient A une matrice carrée de taille m, B une matrice de taille (m, n) *et D une matrice carrée de taille n. Alors*

$$\begin{vmatrix} A & B \\ 0 & D \end{vmatrix} = |A|\,|D|\,.$$

Déterminant d'une matrice par blocs

Exercice 3

Soient A une matrice carrée de taille m, C une matrice de taille (n, m) et D une matrice carrée de taille n. Montrer que

$$\begin{vmatrix} A & 0 \\ C & D \end{vmatrix} = |A|\,|D|\,.$$

Déterminant d'une matrice par blocs

Théorème 4

Soient A une matrice carrée de taille m, B une matrice de taille (m, n) *, C une matrice de taille* (n, m) *et D une matrice carrée de taille n. Alors*

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{cases} |A| \left|D - CA^{-1}B\right| & \text{si } A \text{ est inversible ;} \\ |D| \left|A - BD^{-1}C\right| & \text{si } D \text{ est inversible.} \end{cases}$$

Actualisations de rang 1 et formule de Sherman–Morrison

Théorème 5

Soient A une matrice carrée inversible de taille n et u, v *deux vecteurs colonne de taille* $(n, 1)$. *Alors*

$$\left|A + u\,{}^t v\right| = |A|\,(1 + {}^t v A^{-1} u),$$

et donc $A + u\,{}^t v$ *est inversible si et seulement si* $1 + {}^t v A^{-1} u \neq 0$.
Si $1 + {}^t v A^{-1} u \neq 0$, *l'inverse de* $A + u\,{}^t v$ *est obtenu par la correction de celui de A comme suit :*

$$\left(A + u^t v\right)^{-1} = A^{-1} - \frac{1}{1 +^t vA^{-1} u} A^{-1}\left(u^t v\right) A^{-1}.$$

Produit scalaire

Définition 5

Un produit scalaire sur E est une application $\langle .,. \rangle : E \times E \longrightarrow \mathbb{R}$ satisfaisant les trois propriétés suivantes :

1. $\langle x, y \rangle = \langle y, x \rangle$ *(symétrie).*
2. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ *(linéarité).*
3. $\langle x, x \rangle \geq 0$ *et* $\langle x, x \rangle = 0$ *si et seulement si* $x = 0$.

Norme

Définition 6

Une norme sur E est une application $\|.\| : E \longrightarrow \mathbb{R}^+$ satisfaisant les trois propriétés suivantes :

1. $\|x\| = 0$ *si et seulement si* $x = 0$.
2. $\|\lambda x\| = |\lambda| \, \|x\|$, $\lambda \in \mathbb{R}$.
3. $\|x + y\| \leq \|x\| + \|y\|$ *(inégalité triangulaire).*

Produit scalaire canonique

Définition 7

Le produit scalaire canonique sur $\mathbb{R}^n$ *est défini pour* $x = {}^t(x_1, ..., x_n)$, $y = {}^t(y_1, ..., y_n)$ *par*

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

La norme issue de ce produit scalaire est la norme euclidienne donnée par :

$$\|x\| = \Big(\sum_{i=1}^{n} x_i^2 \Big)^{\frac{1}{2}}.$$

Produit scalaire élliptique

Définition 8

Le produit scalaire élliptique (ou M-produit scalaire) sur $\mathbb{R}^n$ est défini par

$$\langle x, y \rangle_M = {}^t x M y,$$

où M est une matrice carrée de taille n symétrique définie positive (Chapitre 2) générant ce produit scalaire. La norme issue de ce produit scalaire est dite norme élliptique (ou M-norme) et est donnée par :

$$\|x\|_M = \left({}^t x M x \right)^{\frac{1}{2}}.$$

Inégalité de Cauchy Schwarz

Théorème 6

Soit E un espace euclidien. Pour $x, y \in E$ on a

$$|\langle x, y\rangle| \leq \|x\| \, . \, \|y\| \, .$$

Il y a égalité si et seulement si x et y sont linéairement dépendants.

Produit scalaire matriciel canonique

Définition 9

Le produit scalaire canonique sur $\mathcal{M}_{m,n}(\mathbb{R})$ est défini par :

$$\langle A, B\rangle := \langle vec(A), vec(B)\rangle = \sum_{i,j} a_{ij} b_{ij} = \sum_{j} {}^t a_{*j} b_{*j}$$

$$= \sum_{i} {}^t a_{i*} b_{i*} = tr({}^t AB),$$

où l'on a écrit $B = (b_{ij})$ de la même manière par concaténation de ses vecteurs colonnes (resp. lignes).

Norme de Frobenius

Définition 10

La norme de Frobenius (ou norme matricielle standard) est la norme issue du produit scalaire canonique sur $\mathcal{M}_{m,n}(\mathbb{R})$ *, i.e., la norme donnée par*

$$\|A\|^2 = \sum_{i,j} a_{ij}^2 = \sum_j \|a_{*j}\|^2 = \sum_i \|a_{i*}\|^2 = tr({}^tAA).$$

Norme de Frobenius

Proposition 4

1. *la norme de Frobenius est compatible avec la norme (vectorielle) euclidienne dans le sens où l'on a l'inégalité*

$$\|Ax\| \leq \|A\| \, \|x\| \, , \; A \in \mathcal{M}_{m,n}(\mathbb{R}) \, , \; x \in \mathbb{R}^n.$$

2. *la norme de Frobenius est sous-multiplicative, i.e.*

$$\|AB\| \leq \|A\| \, \|B\| \quad \text{si } A \text{ et } B \text{ sont compatibles pour la multiplication}$$

Norme d'opérateur

Définition 11

La norme d'opérateur d'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ est donnée par

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\| \leq 1} \|Ax\| = \max_{\|x\|=1} \|Ax\| .$$

Dans ce cas, la norme d'opérateur sur $\mathcal{M}_{m,n}(\mathbb{R})$ est dite induite par la norme vectorielle.

Norme d'opérateur

Exemple 1

1. *la norme* $\|.\|_p$, $1 \leq p \leq \infty$, *pour dénoter la norme d'opérateur induite par la norme vectorielle* $\|.\|_p$.
2. *La norme matricielle* $\|.\|_2$ *est reconnue souvent pour la norme spectrale et est donnée par :*

$$\|A\|_2 = \sigma_1,$$

où σ_1 *est la plus grande valeur singulière de A.*

3. *Le cas où A est symétrique, sa norme spectrale n'est autre que son rayon spectral :*

$$\|A\|_2 = \rho(A) := \max\{|\lambda_i| : \lambda_i \text{ est une valeur propre de } A\}.$$

pptés d'invariance des normes matricielles

Proposition 5

Soit A une matrice de taille (m, n) . Alors les normes de Frobenius et spéctrale sont

1. *invariantes par passage à la transposée, i.e.* $\|A\| = \| \ {}^tA \ \|$.
2. *orthogonalement invariantes, i.e.,* $\|A\| = \| \ UAV \ \|$, U, V *étant deux matrices orthogonales de tailles m et n respectivement.*

pptés d'invariance des normes matricielles

Exercice 4

Soit $A = (a_{ij})$ *une matrice de taille* (m, n). *Montrer que*

1. $\|A\|_1 = \max_j \sum_i |a_{ij}|$ *(la plus grande* 1*-norme de ses vecteurs colonnes).*
2. $\|A\|_\infty = \max_i \sum_j |a_{ij}|$ *(la plus grande* 1*-norme de ses vecteurs lignes).*
3. *En déduire que* $\| \, {}^tA \, \|_\infty = \|A\|_1$.

Formule de Sherman-Woodbury-Morrison

Exercice 5

1. *Calculer les inverses des matrices par blocs :*
   $\begin{bmatrix} I_{n_1} & 0 \\ X & I_{n_2} \end{bmatrix}$; $\begin{bmatrix} I_{n_1} & Y \\ 0 & I_{n_2} \end{bmatrix}$.
2. *Soit la matrice par blocs* $R = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ *de taille n telles que A et D sont de tailles* n_1 *est* n_2 *resp. Définissons le complément de Schur de A, lorsque celle-ci est inversible, comme étant* $F = D - CA^{-1}B$ *et supposons que A et F sont inversibles. Donner la décomposition LDU par blocs de R :*
$$R = \begin{bmatrix} I_{n_1} & 0 \\ X & I_{n_2} \end{bmatrix} \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} \begin{bmatrix} I_{n_1} & Y \\ 0 & I_{n_2} \end{bmatrix}.$$

Formule de Sherman-Woodbury-Morrison

Exercice 4

1. *En déduire que l'inverse de R est donné par :*

$$R^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BF^{-1}CA^{-1} & -A^{-1}BF^{-1} \\ -F^{-1}CA^{-1} & F^{-1} \end{bmatrix}.$$

2. *Définissons de même le complément de Schur de D, lorsque celle-ci est inversible, comme étant $G = A - BD^{-1}C$ et supposons que D et G sont inversibles. Montrer que*

$$R^{-1} = \begin{bmatrix} G^{-1} & -G^{-1}D^{-1}C \\ -BD^{-1}G^{-1} & D^{-1} + D^{-1}CG^{-1}BD^{-1} \end{bmatrix}.$$

Formule de Sherman-Woodbury-Morrison

Exercice 4

1. *En déduire la formule de Sherman-Woodbury-Morrison lorsque A et D sont inversibles :*

$$\left(A - BD^{-1}C\right)^{-1} = A^{-1} + A^{-1}B\left(D - CA^{-1}B\right)^{-1}CA^{-1}.$$

Orthogonalité dans un espace euclidien

Définition 12

Soit $(E, \langle ., . \rangle)$ un espace euclidien.

1. *L'angle entre deux vecteurs non nuls $x, y \in E$ est défini par*
$$(x \overset{\wedge}{,} y) = \arccos \frac{\langle x, y \rangle}{\|x\| \, \|y\|}.$$
2. *Deux vecteurs $x, y \in E$ sont dits orthogonaux, et on note $x \perp y$, si $\langle x, y \rangle = 0$ (i.e. $(x \overset{\wedge}{,} y) = \frac{\pi}{2}$ si x et y sont non nuls).*

Orthogonalité dans un espace euclidien

Définition 11

1. *Une famille de vecteurs* $X = (x_i)_{i \in I}$ *est dite orthogonale si ses vecteurs* x_i *sont deux à deux orthogonaux. Si de plus, les vecteurs* x_i *sont unitaires, i.e.* $\|x_i\| = 1$ *pour tout* i*, alors* X *est dite orthonormale. On a donc dans ce cas* $\langle x_i, x_j \rangle = \delta_{ij}$*, où* $\delta_{ij} = 0$ *si* $i \neq j$ *et* $\delta_{ij} = 1$ *si* $i = j$ *(le delta de Kronecker).*

Orthogonalité dans un espace euclidien

Théorème 7 (Identité de Pythagore)

Soit E un espace euclidien. Pour deux vecteurs orthogonaux $x, y \in E$, on a

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 .$$

Plus généralement, si $(x_1, x_2, ..., x_p)$ est une famille orthogonale de E, alors

$$\left\|\sum_{i=1}^{p} x_i\right\|^2 = \sum_{i=1}^{p} \|x_i\|^2 .$$

Orthogonalité dans un espace euclidien

Proposition 6

Toute famille orthogonale $(x_1, x_2, ..., x_p)$ *de vecteurs non nuls d'un espace euclidien* E *est libre. Il en est de même, en particulier, pour toute famille orthonormale de* E.

Projection orthogonale sur une droite vectorielle

Proposition 7

Soit $a \in \mathbb{R}^n$ un vecteur non nul. Alors

1. *La matrice de projection orthogonle sur a est donnée par*

$$P_a = \frac{a\,{}^t a}{{}^t a.a}. \tag{2}$$

2. *P_a est idempotente, symétrique et de rang égal à* 1.

Théorème 8 (Orthogonalisation de Gram-Schmidt)

Soit $(x_1, x_2, ..., x_p)$ *une famille libre d'un espace euclidien* E*. Soient les deux familles* $(y_1, y_2, ..., y_p)$ *et* $(e_1, e_2, ..., e_p)$ *construites comme suit :*

$$\begin{aligned} y_1 &= x_1 ; \quad e_1 = \frac{y_1}{\|y_1\|}, \\ y_2 &= x_2 - proj_{y_1}(x_2) ; \quad e_2 = \frac{y_2}{\|y_2\|}, \\ &\vdots \\ y_p &= x_p - \sum_{i=1}^{p-1} proj_{y_i}(x_p) ; \quad e_p = \frac{y_p}{\|y_p\|}, \end{aligned}$$

Théorème 8 (Orthogonalisation de Gram-Schmidt)

Alors $(y_1, y_2, ..., y_p)$ *et* $(e_1, e_2, ..., e_p)$ *sont respectivement une famille orthogonale et une famille orthonormale satisfaisant*

$$Vect\,(x_1, x_2, ..., x_p) = Vect\,(y_1, y_2, ..., y_p) = Vect\,(e_1, e_2, ..., e_p)\,.$$

Théorème 9

Soit $(E, \langle ., . \rangle)$ *un espace euclidien (de dimension* $n \geq 1$*). Alors*

1. E *admet une base orthonormale* $(e_1, e_2, ..., e_n)$.
2. *Tout vecteur* $x \in E$ *s'ecrit* $x = \sum\limits_{i=1}^{n} \langle x, e_i \rangle \, e_i$.
   Cette dernière écriture étant le développement de Fourier de x *et les scalaires* $\langle x, e_i \rangle$ *sont ses coefficients de Fourier.*
3. *La norme de* $x \in E$ *est donnée par l'identité de Parseval :*
   $$\|x\|^2 = \sum_{i=1}^{n} |\langle x, e_i \rangle|^2 .$$
4. *Toute famille orthonormale de* E *peut être complétée en une base orthonormale de* E.

Définition 13

Soit X un sous ensemble non vide d'un espace euclidien E. L'ensemble orthogonal de X, noté $X^\perp$, est défini par

$$X^\perp := \{y \in E : \langle y, x \rangle = 0 \ \forall x \in X\}.$$

Proposition 8

Soit X un sous ensemble non vide d'un espace euclidien E. Alors

1. *Si $y \in X^\perp$ alors $y \in Vect\,(X)^\perp$.*
2. *$X^\perp$ est un sous espace vectoriel de E.*
3. *$Vect\,(X) \subset X^{\perp\perp} := \left(X^\perp\right)^\perp$.*

L'orthogonal d'un sous espace vectoriel

Théorème 10

Pour tout sous espace vectoriel F d'un espace euclidien $(E, \langle ., . \rangle)$, l'orthogonal $F^\perp$ en est un supplémentaire, i.e.,

$$E = F \oplus F^\perp.$$

Corollaire 2

Soit F un sous espace vectoriel d'un espace euclidien E. Alors $F^{\perp\perp} = F$.

Projection orthogonale

Définition 14

Soit F un sous espace vectoriel de $\mathbb{R}^n$. La projection orthogonale d'un vecteur $x \in \mathbb{R}^n$ sur F est la projection de x sur F parallèlement à $F^{\perp}$. La matrice de cette projection est dite matrice de projection orthogonale sur F.

Théorème 11

Soient F un sous espace vectoriel de $\mathbb{R}^n$ et $x \in \mathbb{R}^n$. La projection orthogonale de x sur F est l'unique vecteur $p(x)$ de F qui réalise

$$\|x - p(x)\| = \min_{u \in F} \|x - u\|.$$

Les quatre sous espaces vectoriels fondamentaux d'une matrice

Soit A une matrice de taille (m, n) et de rang r. Les quatre SEVF associés à la matrice A sont les images et noyaux de A et de sa transposée :

- ImA un sous espace vectoriel de $\mathbb{R}^m$ de dimension r.
- $KerA$ un sous espace vectoriel de $\mathbb{R}^n$ de dimension $n - r$.
- Im ${}^tA$ un sous espace vectoriel de $\mathbb{R}^n$ de dimension r.
- $Ker\ {}^tA$ un sous espace vectoriel de $\mathbb{R}^m$ de dimension $m - r$.

Détermination des bases des 4 SEVF

Proposition 9

Soit A une matrice de taille (m, n) et de rang r. Alors

1. $Im\,{}^tA = Im\,{}^tR$ *et donc les r vecteurs lignes pivots de R forment une base de $Im\,{}^tA$.*
2. $KerA = KerR$ *et donc les $n - r$ solutions particulières et linéairement indépendantes du système homogène $Rx = 0$ forment une base de $KerA$.*

Détermination des bases des 4 SEVF

On utilise

Exercice 6

Soient A et B deux matrices de taille (m, n) *et* (n, p) *respectivement.*

1. *Montrer que* $Im(AB) \subset ImA$ *et qu'il y a égalité si B est une matrice carrée inversible.*
2. *Montrer que* $KerB \subset Ker\,(AB)$ *et qu'il y a égalité si A est une matrice carrée inversible.*

Détermination des bases des 4 SEVF

Proposition 10

Soient A et B deux matrices ayant le même nombre de colonnes. Alors

1. *si* $KerA = KerB$, *alors une famille de vecteurs colonnes de A est libre si et seulement si la famille de vecteurs colonnes de B dans les positions correspondantes est libre ;*
2. *une famille de vecteurs colonnes de A est libre si et seulement si la famille de vecteurs colonnes de R dans les positions correspondantes est libre. Donc, les vecteurs colonnes pivots de* A, *i.e., les vecteurs colonnes occupant les positions correspondantes des vecteurs colonnes de R contenant les pivots, forment une base de ImA.*

Détermination des bases des 4 SEVF

Proposition 11

Ecrivons $G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}$ *avec* G_1 *est de taille* (r, m) *(et donc* G_2 *est de taille* $(m-r, m)$*). Alors,* $\operatorname{Ker} {}^t A = \operatorname{Im} {}^t G_2$ *et donc les* $m - r$ *vecteurs lignes de* G_2 *forment une base de* $\operatorname{Ker} {}^t A$*.*

Le théorème fondamental de l'algèbre linéaire

Théorème 12

Soit A une matrice de taille (m, n) et de rang r. Alors

1. $(ImA)^{\perp} = Ker\ {}^{t}A$ *et on a* $\mathbb{R}^m = ImA \oplus Ker\ {}^{t}A$;
2. $(KerA)^{\perp} = Im\ {}^{t}A$ *et on a* $\mathbb{R}^n = KerA \oplus Im\ {}^{t}A$.

Le théorème fondamental de l'algèbre linéaire

Exercice 7 (Alternative de Fredholm)

Soient A une matrice de taille (m, n) et $b \in \mathbb{R}^m$. Montrer qu'exactement un des deux problèmes suivants a une solution :

1. $Ax = b$, $x \in \mathbb{R}^n$.
2. ${}^t y.A = 0$, $y \in \mathbb{R}^m$ *avec* ${}^t y.b \neq 0$.

Projection orthogonale et l'equation normale

Exercice 8

Soit A une matrice de taille (n, r) et de rang plein. Vérifier que

1. *P_A est une matrice idempotente, symétrique et de rang égal à r.*
2. *$P_A b = b$ si $b \in ImA$ et $P_A b = 0$ si $b \in (ImA)^{\perp}$.*

Projection orthogonale et l'equation normale

Théorème 13

Soient A une matrice de taille (n, r) et de rang plein. Alors

1. *La projection orthogonale de $b \in \mathbb{R}^n$ sur ImA est donnée par $A\hat{x}$ où $\hat{x} \in \mathbb{R}^r$ est l'unique solution de l'équation normale :*
$$ {}^tAA\hat{x} = {}^tAb. $$
2. *La matrice de projection orthogonale sur ImA est donnée par*
$$ P_A = A({}^tAA)^{-1}\,{}^tA. $$
3. *P_A est une matrice idempotente, symétrique et de rang égal à r.*

Projection orthogonale et l'equation normale

Définition 15

Soient A une matrice de taille (n, r) et de rang plein et $b \in \mathbb{R}^n$. L'unique solution de l'équation normale :

$$ {}^tAA\hat{x} = {}^tAb. $$

réalise bien

$$ \min_{x \in \mathbb{R}^r} \|Ax - b\| = \|A\hat{x} - b\|. $$

Elle est donc dite la solution au sens des moindres carrés du système linéaire $Ax = b$.

Projection orthogonale et l'equation normale

Proposition 12

Soient A et B deux matrices de taille (n, r) *et de rang plein telles que ImA =ImB. Alors* $P_A = P_B$.

Projection orthogonale et l'equation normale

Exercice 9 (Contrôle final, automne 2018)

Soient A une matrice réelle de taille (m, n) *(*$m > n$*) et de rang plein,* $Ax = b$ *un système linéaire surdéterminé et* $\hat{x}$ *sa solution au sens des moindres carrés. On actualise le système* $Ax = b$ *par l'ajout de k équations représentées par le système linéaire* $Vx = c$ *où* V *est une matrice de taille* (k, n) *et* $c \in \mathbb{R}^k$*. Montrer que la solution actualisée au sens des moindres carrés est donnée par*

$$\hat{x}_a = \hat{x} + K(c - V\hat{x})$$

où $K = B^{-1}\,{}^tV(I_k + VB^{-1}\,{}^tV)^{-1}$ *(la matrice* $I_k + VB^{-1}\,{}^tV$ *étant inversible) et* $B = {}^tAA$*.*
(Utiliser la formule de Sherman-Woodbury-Morrison).

Projection orthogonale et l'equation normale

Exercice 10 (Contrôle final, automne 2017)

1. *Soient A et B deux matrices de tailles (m, p) et (m, q) respectivement. Montrer que $ImA \perp ImB$ si et seulement si $\left({}^tA\right) B = 0$ et $\left({}^tB\right) A = 0$.*
2. *Supposons que la matrice $C = [A, B]$ est de rang plein. Montrer que $P_C = P_A + P_B$.*

Exercice 11 (Formule matricielle d'un projecteur)

Soient $\mathbb{R}^n = F \oplus G$, $(a_1, ..., a_r)$ *une base de* F $(r < n)$ *et* $(b_1, ..., b_{n-r})$ *une base de* G*. Soient les matrices* $A = [a_1, ..., a_r]$, $B = [b_1, ..., b_{n-r}]$, $S = [A, B]$ *et* P *est la matrice de projection sur* F *parallèlement à* G*.*

1. *Calculer* PS *et en déduire que* $P = AC$ *où* C *est une matrice à déterminer.*
2. *Soient* $(v_1, ..., v_r)$ *une base de* $G^{\perp}$ *et* $V = [v_1, ..., v_r]$ *.*
   1. *Montrer que* ${}^tV = {}^tV\ AC$.
   2. *En déduire que* $Im({}^tVA) = Im({}^tV)$ *et que* ${}^tVA$ *est inversible.*
3. *En déduire que la matrice* P *est donnée par la formule :*

$$P = A({}^tVA)^{-1}\ {}^tV.$$

4. *Commenter ce résultat*

Projection orthogonale et l'equation normale

Exercice 12 (Méthode des moindres carrés - approche différentielle)

Soient $A = (a_{ij})$ *une matrice carrée de taille n et* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, *définie par*

$$f(x) = \langle Ax, x \rangle = {}^t xAx.$$

1. *Donner* $f(x)$ *en fonction des* x_i *et* a_{ij}, *où* $x = {}^t(x_1, ..., x_n)$. *En déduire l'expression de la dérivée partielle* $\frac{\partial f}{\partial x_p}$, $p = 1, 2, ..., n$.

Projection orthogonale et l'equation normale

Exercice 12 (Méthode des moindres carrés - approche différentielle)

1. *Soit* ∇f *le gradient de* f *défini par* $\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}$, $x \in \mathbb{R}^n$. *Montrer que*

$$\nabla f(x) = (A + {}^tA)\, x, \quad x \in \mathbb{R}^n.$$

Projection orthogonale et l'equation normale

Exercice 12 (Méthode des moindres carrés - approche différentielle)

1. *Soient B une matrice de taille (m, n) , $x \in \mathbb{R}^n$ et $b \in \mathbb{R}^m$. Montrer que*

$$\|Bx - b\|^2 = {}^t x\, {}^t BBx \; - 2\, {}^t x\, {}^t Bb + {}^t bb.$$

2. *En déduire que si $\|Bx - b\|^2$ est minimum alors x est solution de l'équation normale associée au système linéaire $Bx = b$.*

Application : ajustement affine par la méthode des moindres

- Soit $\mathcal{N}$ un nuage de points (x_i, y_i) résultat de n observations x_i et y_i, $i = 1, 2, ..., n$, de deux variables statistiques quantitatives X et Y respectivement, telles que les réels x_i ne sont pas tous égaux.
- On cherche à trouver la droite D ajustant le mieux possible le nuage de points $\mathcal{N}$.
- La méthode des moindres carrés : trouver la droite D, dite droite des moindres carrés, parmi les droites d'équation $Y = a + bX$ où a, b sont des constantes réelles, telle que D minimise la somme des carrés des résidus (ou erreurs)

$$e = \sum_{i=1}^{n} (y_i - (a + bx_i))^2 = \|\varepsilon\|^2, \text{ où } \varepsilon = (\varepsilon_1, ..., \varepsilon_n),$$

$\varepsilon_i = y_i - (a + bx_i)$, est le vecteur des résidus

Exercice 13 (Contrôle continu, automne 2017)

Soit le nuage de points (x_i, y_i) *comme précédemment. Posons* $x = {}^t(x_1, x_2, ..., x_n)$ *et* $y = {}^t(y_1, y_2, ..., y_n)$.

1. *Donner les écritures matricielles, en fonction de* x *et d'un vecteur colonne* e *à déterminer, de la moyenne arithmétique* $\overline{X}$ *et de la variance* $V(X)$ *des* x_i, $i = 1, 2, ..., n$.
2. *Donner l'écriture matricielle, en fonction de* x *et* y, *de la covariance* $Cov(X, Y)$ *des deux variables* X *et* Y.
3. *Soit* $\varepsilon(u) = y - Au$ *le vecteur des résidus et posons* $\varepsilon(u) = {}^t(\varepsilon_1(u), \varepsilon_2(u), ..., \varepsilon_n(u))$. *Montrer que* $\sum_{i=1}^{n} \varepsilon_i\left(\hat{u}\right) = 0$, *où* $\hat{u}$ *est l'unique solution de l'équation normale* (E) *associée au système linéaire* $Au = y$.

Exercice 11

1. *Écrire la matrice A à l'aide des deux vecteurs colonnes e et x et montrer que l'équation normale (E) s'écrit :*

$$\left\{ \begin{array}{r} na + \left({}^t e.x\right) b = {}^t e.y \\ \left({}^t e.x\right) a + \left({}^t x.x\right) b = {}^t x.y \end{array} \right.$$

2. *En déduire que la droite D a pour équation :*

$$Y = \frac{Cov\,(X, Y)}{V\,(X)} \left(X - \overline{X}\right) + \overline{Y}.$$

Matrices orthogonales

Proposition 13

Pour une matrice carrée Q de taille n, les assertions suivantes sont équivalentes :

1. *les vecteurs colonne de Q forment une famille orthonormale de $\mathbb{R}^n$.*
2. ${}^tQQ = Q\,{}^tQ = I_n$ *, i.e., Q est inversible et* $Q^{-1} = {}^tQ$.

Définition 16

Une matrice carrée A de taille n est dite matrice orthogonale s'elle vérifie l'une des deux assertions de la proposition précédente.

Matrices orthogonales

Exemple 2

1. *Matrices de rotation dans* $\mathbb{R}^2$:

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

On a $R(\theta)^{-1} = {}^tR(\theta) = R(-\theta)$.

2. *Matrice de permutation :*

$$Q = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Matrices orthogonales

Proposition 14

Pour une matrice orthogonale Q de taille n, on a

1. $\langle Qx, Qy \rangle = \langle x, y \rangle$, $x, y \in \mathbb{R}^n$;
2. *en particulier,* $\|Qx\| = \|x\|$, $x \in \mathbb{R}^n$.

Décomposition *QR*

Théorème 14 (Décomposition *QR*)

Soit A une matrice de taille (n, p) et de rang plein. Alors A admet une factorisation $A = QR$ avec Q est une matrice de taille (n, p) dont les vecteurs colonne forment une base orthonormale de $\mathrm{Im}A$ et R est une matrice triangulaire suppérieure de taille p dont les éléments diagonaux sont tous strictement positifs.

Décomposition *QR*

Exercice 14 (Unicité de la décomposition *QR*)

1. *Montrer que pour toute matrice R triangulaire suppérieure dont les éléments diagonaux sont tous strictement positifs, si ${}^tRR = I$ alors $R = I$.*
2. *En déduire que la décomposition QR du Théorème 14 est unique.*

Décomposition QR

Exercice 15 (Décomposition QR- autre version)

Soit A une matrice de taille (n, p) et de rang plein. Soit $A = Q_1 R_1$ la décomposition du Théorème 14.

1. *MATLAB est programmé pour donner une décomposition $A = QR$ telle que $Q = [Q_1, Q_2]$ est une matrice orthogonale de taille n et R est une matrice de taille (n, p). Montrer qu'une telle décomposition existe.*
2. *Quel sous espace véctoriel parmi les quatre sous espaces véctoriels fondamentaux de A les vecteurs colonnes de Q_2 en forment une base ?*
3. *La décomposition $A = QR$ est-elle unique ?*

Décomposition *QR*

Exercice 16

Soit le système linéaire $Ax = b$, $b = {}^t(b_1, ..., b_m)$, *à résoudre au sens des moindres carrés. On multiplie les m équations de ce système par des poids* $w_1, ..., w_m > 0$.
Montrer que l'équation normale s'écrit

$$ {}^tACAx = {}^tACb $$

où C est une matrice à déterminer.

Définition 17

1. *Soit A une matrice (réelle) carrée de taille n. Un scalaire $\lambda \in \mathbb{C}$ est dit une valeur propre de A s'il existe $x \in \mathbb{C}^n \setminus \{0\}$ solution de l'équation $Ax = \lambda x$. Le vecteur x est dit un vecteur propre de A associé à la valeur propre λ.*
2. *Le sous espace vectoriel $Ker(A - \lambda I_n)$ de $\mathbb{C}^n$ est dit le sous espace propre de A associé à λ. La dimension de $Ker(A - \lambda I_n)$ est dite la multiplicité géométrique de la valeur propre λ et est notée $MG_A(\lambda)$.*
3. *Le polynôme d'indéterminée X*

$$\chi_A(X) = |XI_n - A| = (X - \lambda_1)(X - \lambda_2)\dots(X - \lambda_n)$$

est dit le polynôme caractéristique de A.

Produit hermitien

Définition 18

Le produit scalaire (ou produit hermitien) canonique sur le $\mathbb{C}$-espace vectoriel $\mathbb{C}^n$ est défini par

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

où pour $x = {}^t(x_1, ..., x_n) \in \mathbb{C}^n$, $x^* := {}^t\bar{x} = \overline{{}^t x} = (\bar{x}_1, ..., \bar{x}_n)$.
La norme issue de ce produit scalaire est donnée par

$$\|x\| = (x^* x)^{\frac{1}{2}} = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}}$$

Définition 19

1. *Le nombre de fois que se répéte une valeur propre λ en tant que racine du polynôme caractéristique de A est dit la multiplicité algébrique de λ et est notée $MA_A(\lambda)$.*
2. *Une matrice carrée A de taille n est dite diagonalisable s'elle est semblable à une matrice diagonale D de taille n à coefficients dans $\mathbb{C}$ via une matrice inversible P de taille n à coefficients dans $\mathbb{C}$, i.e., $A = PDP^{-1}$.*

Théorème 15

Soit A une matrice carrée de taille n. Alors, les assertions suivantes sont équivalentes :

1. *A est diagonalisable ;*
2. *toute valeur propre de A est régulière ;*
3. *si* $\sigma(A) = \{\lambda_1, \lambda_2, ..., \lambda_k\}$ *est l'ensemble des valeurs propres de A alors*

   $$\mathbb{C}^n = Ker\,(A - \lambda_1 I_n) \oplus Ker\,(A - \lambda_2 I_n) \oplus ... \oplus Ker\,(A - \lambda_k I_n)\,;$$

4. *il existe une famille libre* $(x_1, ..., x_n)$ *de* $\mathbb{C}^n$ *constituée de vecteurs propres de A.*

Matrices réelles symétriques

Proposition 15

Soit A une matrice réelle symétrique de taille n. Alors

1. *les valeurs propres de A sont réelles ;*
2. *deux vecteurs propres réels de A associés à deux valeurs propres distinctes sont orthogonaux.*

Matrices réelles symétriques

Théorème 16 (Le théorème spectral)

Soit A une matrice réelle symétrique de taille n. Alors, il existe une base orthonormale $(p_1, ..., p_n)$ de $\mathbb{R}^n$ formée de vecteurs propres de A associés respectivement à ses valeurs propres $\lambda_1, ..., \lambda_n$. Les équations

$$Ap_i = \lambda_i p_i \tag{3}$$

s'écrivent matriciellement

$$AP = PD \text{ soit } A = PD\,{}^tP \tag{4}$$

où $P = [p_1, ..., p_n]$ est une matrice orthogonale de taille n, dite matrice de changement de bases, et $D = diag\,(\lambda_1, ..., \lambda_n)$.

Le quotient de Rayleigh

Pour une matrice carrée A de taille n et $x \in \mathbb{R}^n \setminus \{0\}$, on considère le problème de minimisation :

$$arg \min_{\alpha \in \mathbb{R}} \|Ax - \alpha x\| . \tag{5}$$

Proposition 16

La solution du problème (5) *est donné par le quotient*

$$R_A(x) = \frac{{}^t x A x}{{}^t x x}.$$

Le quotient $R_A(x)$, $x \in \mathbb{R}^n \setminus \{0\}$, *est appelé le quotient de Rayleigh.*

Le quotient de Rayleigh

Théorème 17 (Extremums du quotient de Rayleigh)

Soit A une matrice réelle symétrique de taille n. Alors le quotient de Rayleigh satisfait :

1. $\max_{x\neq 0} \frac{{}^t xAx}{{}^t xx} = \lambda_1$, λ_1 *étant la plus grande valeur propre de A et le maximum est atteint en x_1 un vecteur propre de A associé à λ_1.*
2. $\min_{x\neq 0} \frac{{}^t xAx}{{}^t xx} = \lambda_n$, λ_n *étant la plus petite valeur propre de A et le minimum est atteint en x_n un vecteur propre de A associé à λ_n.*

Le quotient de Rayleigh

Exercice 17 (Extremums du quotient de Rayleigh - approche analytique)

Soient A une matrice carrée de taille n et $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par $f(x) = {}^t xAx$.

1. *Montrer que f admet un maximum et un minimum sur la sphère $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$.*
2. *En déduire que le quotient de Rayleigh $R_A(x)$, $x \in \mathbb{R}^n \setminus \{0\}$, admet un maximum $R_A(x_1)$ et un minimum $R_A(x_2)$.*
3. *Supposons que A est symétrique. Calculer $\frac{d}{dx} R_A(x)$, pour $x \in \mathbb{R}^n \setminus \{0\}$, et en déduire que le maximum et le minimum de $R_A(x)$ sont respectivement la plus grande et la plus petite valeur propre de A et que x_1 et x_2 en sont*

Matrices (semi) définies positives

Définition 20

Soit A une matrice symétrique de taille n.

1. *A est dite semi-définie positive si ${}^t xAx \geq 0$ pour tout $x \in \mathbb{R}^n$.*
2. *A est dite définie positive si ${}^t xAx > 0$ pour tout $x \in \mathbb{R}^n \setminus \{0\}$.*

Matrices (semi) définies positives

Proposition 17

Soit la matrice diagonale $D = diag\,(d_1, ..., d_n)$ *. Alors*

1. D *est semi-définie positive ssi* $d_i \geq 0$ *pour tout* i.
2. D *est définie positive ssi* $d_i > 0$ *pour tout* i.

Proposition 18

Soient A *une matrice symétrique de taille* n *et* P *une matrice de taille* (n, m) .

1. *Si* A *est semi-définie positive alors* ${}^tPAP$ *l'est aussi ;*
2. *si* A *est définie positive et* P *est de rang plein (en particulier, si* P *est une matrice carrée inversible), alors* ${}^tPAP$ *est définie positive.*

Matrices (semi) définies positives

Théorème 18

Soit A une matrice symétrique de taille n.

1. *A est semi-définie positive si et seulement si toutes les valeurs propres λ_i de A sont positives.*
2. *A est définie positive si et seulement si toutes les valeurs propres λ_i de A sont strictement positives.*

Matrices (semi) définies positives

Corollaire 3

Soit A une matrice symétrique de taille n. Alors

1. *si A est définie positive alors* $|A| > 0$ *(et donc A est inversible) et* A^{-1} *est aussi définie positive.*
2. *si A est semi-définie positive et inversible alors A est définie positive.*

Matrices (semi) définies positives

Proposition 19

Pour toute matrice A de taille (m, n) on a

1. ${}^{t}AA$ *est semi-définie positive ;*
2. *si A est de rang plein alors ${}^{t}AA$ est définie positive.*

Matrices (semi) définies positives

Théorème 19

Soit A une matrice symétrique de taille n. Alors

1. *A est semi-définie positive si et seulement s'il existe une matrice carrée B de taille n telle que $A = {}^tBB$;*
2. *A est définie positive si et seulement s'il existe une matrice carrée inversible B de taille n telle que $A = {}^tBB$.*

Matrices (semi) définies positives

Exercice 18

Soit A une matrice de taille (m, n). Considérons le problème

$$\begin{cases} \text{maximiser } f(x) = \|Ax\|^2, \ x \in \mathbb{R}^n \\ \text{sous la contrainte } g(x) = {}^t xx = 1. \end{cases} \tag{6}$$

1. *Soit la fonction $h(x, \lambda) = f(x) - \lambda g(x)$, où λ est le multiplicateur de Lagrange. Calculer $\frac{d}{dx} h(x, \lambda)$.*
2. *En déduire que la solution du problème (6) est la plus grande valeur propre σ_1^2 $(\sigma_1 \geq 0)$ de ${}^t AA$, et que*

$$\|A\|_2 = \sigma_1.$$

Matrices (semi) définies positives

Exercice 16

1. *Montrer que si A est symétrique, alors*

$$\|A\|_2 = \rho(A) := \max\{|\lambda_i| : \lambda_i \in \sigma(A)\}.$$

2. *Montrer que si A est une matrice carrée inversible, alors*

$$\left\|A^{-1}\right\|_2 = \frac{1}{\sigma_n},$$

où σ_n^2 *(*$\sigma_n > 0$*) est la plus petite valeur propre de* ${}^tAA$.

Théorème 20 (Décomposition en valeurs singulières)

Toute matrice A de taille (m, n) et de rang $r \geq 1$ admet une décomposition $A = U\Sigma\,{}^tV$, dite décomposition en valeurs singulières, où $U = [u_1, ..., u_m]$, $V = [v_1, ..., v_n]$ sont deux matrices orthogonales de tailles m et n respectivement, $\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$ est une matrice de taille (m, n), et $\Sigma_r = diag\,(\sigma_1, ..., \sigma_r)$, $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0$ sont les racines carrées des r valeurs propres strictement positives de ${}^tAA$ et sont dites les valeurs singulières de A.

Théorème 20 (Décomposition en valeurs singulières)

De plus, les vecteurs $v_1, ..., v_n$ sont dits des vecteurs singuliers à droite de A et satisfont

- *les r premiers vecteurs $v_1, ..., v_r$ forment une base orthonormale de Im ^tA.*
- *les $n - r$ derniers vecteurs $v_{r+1}, ..., v_n$ forment une base orthonormale de KerA.*

Aussi, les vecteurs $u_1, ..., u_m$ sont dits des vecteurs singuliers à gauche de A et satisfont

- *les r premiers vecteurs $u_1, ..., u_r$ forment une base orthonormale de ImA.*
- *les $m - r$ derniers vecteurs $u_{r+1}, ..., u_m$ forment une base orthonormale de Ker ^tA.*

Exercice 19

Soient A une matrice symétrique de taille n et de rang $r \geq 1$, $A = PD\,{}^tP$ *une décomposition spectrale avec P est une matrice orthogonale de taille n et* $D = diag\left(\lambda_1, ..., \lambda_n\right)$, *les* λ_i *étant les valeurs propres de A arrangées de telle façon que* $|\lambda_1| \geq |\lambda_2| \geq ... \geq |\lambda_r| > \lambda_{r+1} = ... = \lambda_n = 0$.

1. *Donner une décomposition en valeurs singulières* $A = U\Sigma\,{}^tV$, *où l'on déterminera les matrices* U, V *en fonction de P et d'une matrice* Δ *à déterminer, et* Σ *en fonction des* λ_i.

Exercice 20

2. *En déduire les valeurs singulières de A en fonction des* λ_i.
3. *Sous quelle condition supplémentaire sur A, ses deux décompositions spéctrale et en valeurs singulières ci-dessus sont les mêmes ?*

Exercice 21

Montrer que la norme d'une matrice A (non nulle) de taille (m, n) est donnée par

$$\|A\| = (\sigma_1^2 + ... + \sigma_r^2)^{\frac{1}{2}}$$

où les σ_i sont les valeurs singulières de A et $r = \mathrm{rg}A$.

Exercice 22

Soit $A = U\Sigma\, {}^tV$ *une décomposition en valeurs singulières d'une matrice* A *de taille* (m, n) *telles que* $U = [u_1, u_2, ..., u_m]$ *et* $V = [v_1, v_2, ..., v_n]$*. Montrer que la famille*

$$\mathcal{F} = \left(u_i\, {}^t v_j\right)_{1 \leq i \leq m,\ 1 \leq j \leq n}$$

est une base orthonormale de $\mathcal{M}_{m,n}(\mathbb{R})$.

Exercice 23

Soit A une matrice de taille (n, m) *(*$m \leq n$*). Montrer que le quotient* $\frac{{}^t xAy}{\|x\|\|y\|}$, $x \in \mathbb{R}^n \setminus \{0\}$ *et* $y \in \mathbb{R}^m \setminus \{0\}$, *admet un maximum et déterminer deux vecteurs* x *et* y *pour lesquels ce maximum est atteint.*

Exercice 24

On se propose dans cet exercice de montrer le résultat de minimisation suivant, utile en analyse factorielle : pour deux matrices réelles A et B de taille (n, p), *on a*

$$\min_{X \in \mathcal{O}_p} \|A - BX\|^2 = \|A - BX_*\|^2$$

où $\mathcal{O}_p$ *est l'ensemble des matrices orthogonales de taille p et* $X_* = U\,{}^tV$, *U et V étant obtenues de la décomposition en valeurs singulières* ${}^tBA = U\Sigma\,{}^tV$.

1. *Montrer que pour tout* $X \in \mathcal{O}_p$, *on a*

$$\|A - BX\|^2 = tr(A\,{}^tA) + tr(B\,{}^tB) - 2tr(X\,{}^t({}^tBA)).$$

Exercice 25

2. *Montrer que pour tout* $X \in \mathcal{O}_p$, $tr(X\,{}^t({}^tBA)) \leq tr\,(\Sigma)$ *et conclure.*
   (ind. les coefficients d'une matrice orthogonale sont majorés par 1*).*

Exercice 26 (théorème d'Eckart-Young)

Pour des raisons de réduction de la dimension de l'espace vectoriel ImZ*, on s'intéresse à la recherche d'un sous-espace vectoriel de* ImZ *de dimension* $k < r = rg(Z)$ *approchant* ImZ *dans un certain sens. Pour cela, on cherche une matrice* Z_k *de même taille que* Z *et de rang* k *telle que*

$$\|Z - Z_k\|_F = \min_{\substack{S \in \mathcal{M}_{n,p}(\mathbb{R}) \\ rg(S) = k}} \|Z - S\|_F \qquad (*)$$

où $\|.\|_F$ *désigne la norme de Frobenius sur l'espace vectoriel* $\mathcal{M}_{n,p}(\mathbb{R})$ *des matrices réelles de taille* (n, p)*.*

Exercice 27

1. *Soit $Z = U\Sigma^t V$ une décomposition en valeurs singulières de Z telles que $U = [u_1, u_2, ..., u_n]$, $V = [v_1, v_2, ..., v_p]$ sont deux matrices orthogonales et $\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{M}_{n,p}(\mathbb{R})$ avec $\Sigma_r = diag(\sigma_1, \sigma_2, ..., \sigma_r)$ et $\mathcal{F} = \left(u_i{}^t v_j\right)_{1\leq i\leq n,\, 1\leq j\leq p}$.*
2. *Soit (α_{ij}) les coordonées d'une matrice arbitraire $S \in \mathcal{M}_{n,p}(\mathbb{R})$ dans la base $\mathcal{F}$. Montrer que*

$$\|Z - S\|_F^2 = \sum_{i=1}^{r} (\sigma_i - \alpha_{ii})^2 + \sum_{r<i\leq \min(n,p)} \alpha_{ii}^2 + \sum_{i\neq j} \alpha_{ij}^2.$$

Exercice 28

1. *En déduire que la matrice Z_k est donnée par la troncature de la décomposition en valeurs singulières $Z = \sum_{i=1}^{r} \sigma_i u_i{}^t v_i$ jusqu'à l'ordre k, i.e., $Z_k = \sum_{i=1}^{k} \sigma_i u_i{}^t v_i$ et que le minimum dans l'égalité $(*)$ vaut $\left(\sum_{i=k+1}^{r} \sigma_i^2\right)^{\frac{1}{2}}$.*

Exercice 29

On se propose dans cet exercice de montrer le résultat de minimisation suivant, utile en analyse factorielle : pour deux matrices réelles A et B de taille (n, p), on a

$$\min_{X \in \mathcal{O}_p} \|A - BX\|^2 = \|A - BX_*\|^2$$

où $\mathcal{O}_p$ est l'ensemble des matrices orthogonales de taille p et $X_* = U\,{}^tV$, U et V étant obtenues de la décomposition en valeurs singulières ${}^tBA = U\Sigma\,{}^tV$.

1. *Montrer que pour tout $X \in \mathcal{O}_p$, on a*

$$\|A - BX\|^2 = tr(A\,{}^tA) + tr(B\,{}^tB) - 2tr(X\,{}^t({}^tBA)).$$

Exercice 30

2. *Montrer que pour tout $X \in \mathcal{O}_p$, $tr(X\,{}^t({}^tBA)) \leq tr\,(\Sigma)$ et conclure.*
   (ind. les coefficients d'une matrice orthogonale sont majorés par 1).