Rappels et compléments d'algèbre linéaire

Résumé

Cette vignette rassemble des notations et rappels d'algèbre linéaire de niveau L. Il introduit les principaux théorèmes d'approximation matricielle par décomposition en valeurs singulières qui sont à la base des méthodes statistique factorielles.

Retour au plan du cours.

1 Notations

Dans tout ce qui suit, E et F sont deux espaces vectoriels réels munis respectivement des bases canoniques $\mathcal{E} = \{\mathbf{e}_j \; ; \; j = 1, \ldots, p\}$ et $\mathcal{F} = \{\mathbf{f}_i \; ; \; i = 1, \ldots, n\}$. On note indifféremment soit un vecteur de E ou de F, un endomorphisme de E, ou une application linéaire de E dans F, soit leurs représentations matricielles dans les bases définies ci-dessus.

2 Matrices

2.1 Notations

La matrice d'ordre $(n \times p)$ associée à une application linéaire de E dans F est décrite par un tableau :

$$\mathbf{A} = \begin{bmatrix} a_1^1 & \ldots & a_1^j & \ldots & a_1^p \\ \vdots & & \vdots & & \vdots \\ a_i^1 & \ldots & a_i^j & \ldots & a_i^p \\ \vdots & & \vdots & & \vdots \\ a_n^1 & \ldots & a_n^j & \ldots & a_n^p \end{bmatrix}.$$

On note par la suite :

$$\begin{aligned} a_i^j &= [\mathbf{A}]_i^j \text{ le terme général de la matrice,} \\ \mathbf{a}_i &= [a_i^1, \ldots, a_i^p]' \text{ un vecteur-ligne mis en colonne,} \\ \mathbf{a}^j &= [a_1^j, \ldots, a_n^j]' \text{ un vecteur-colonne.} \end{aligned}$$

2.1.1 Types de matrices

Une matrice est dite :

- *vecteur-ligne (colonne)* si $n = 1$ $(p = 1)$,
- *vecteur-unité* d'ordre p si elle vaut $\mathbf{1}_p = [1, \ldots, 1]'$,
- *scalaire* si $n = 1$ et $p = 1$,
- *carrée* si $n = p$.

Une matrice carrée est dite :

- *identité* $(\mathbf{I}_p)$ si $a_i^j = \delta_i^j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$,
- *diagonale* si $a_i^j = 0$ lorsque $i \neq j$,
- *symétrique* si $a_i^j = a_j^i, \forall (i,j)$,
- *triangulaire* supérieure (inférieure) si $a_i^j = 0$ lorsque $i > j$ $(i < j)$.

2.1.2 Matrice partitionnée en blocs

Matrices dont les éléments sont eux-mêmes des matrices. Exemple :

$$\mathbf{A}(n \times p) = \begin{bmatrix} \mathbf{A}_1^1(r \times s) & \mathbf{A}_1^2(r \times (p-s)) \\ \mathbf{A}_2^1((n-r) \times s) & \mathbf{A}_2^2((n-r) \times (p-s)) \end{bmatrix}.$$

2.2 Opérations sur les matrices

Somme : $[\mathbf{A} + \mathbf{B}]_i^j = a_i^j + b_i^j$ pour $\mathbf{A}$ et $\mathbf{B}$ de même ordre $(n \times p)$.

Multiplication par un scalaire : $[\alpha \mathbf{A}]_i^j = \alpha a_i^j$ pour $\alpha \in \mathbf{R}$.

Transposition : $[\mathbf{A}']_i^j = a_j^i$, $\mathbf{A}'$ est d'ordre $(p \times n)$.

$$(\mathbf{A}')' = \mathbf{A} \, ; (\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}' \, ; (\mathbf{AB})' = \mathbf{B}'\mathbf{A}' \, ;$$

$$\begin{bmatrix} \mathbf{A}_1^1 & \mathbf{A}_1^2 \\ \mathbf{A}_2^1 & \mathbf{A}_2^2 \end{bmatrix}' = \begin{bmatrix} \mathbf{A}_1^{1'} & \mathbf{A}_2^{1'} \\ \mathbf{A}_1^{2'} & \mathbf{A}_2^{2'} \end{bmatrix}.$$

Produit scalaire élémentaire : $a'b = \sum_{i=1}^n a_i b_i$ où a et b sont des vecteurs-colonnes.

Produit : $[\mathbf{AB}]_i^j = a_i' b^j$ avec $\mathbf{A}_{(n\times p)}, \mathbf{B}_{(p\times q)}$ et $\mathbf{AB}_{(n\times q)}$, et pour des matrices par blocs :

$$\left[\begin{array}{cc} \mathbf{A}_1^1 & \mathbf{A}_1^2 \\ \mathbf{A}_2^1 & \mathbf{A}_2^2 \end{array}\right]\left[\begin{array}{cc} \mathbf{B}_1^1 & \mathbf{B}_1^2 \\ \mathbf{B}_2^1 & \mathbf{B}_2^2 \end{array}\right] = \left[\begin{array}{cc} \mathbf{A}_1^1\mathbf{B}_1^1 + \mathbf{A}_1^2\mathbf{B}_2^1 & \mathbf{A}_1^1\mathbf{B}_1^2 + \mathbf{A}_1^2\mathbf{B}_2^2 \\ \mathbf{A}_2^1\mathbf{B}_1^1 + \mathbf{A}_2^2\mathbf{B}_2^1 & \mathbf{A}_2^1\mathbf{B}_1^2 + \mathbf{A}_2^2\mathbf{B}_2^2 \end{array}\right]$$

sous réserve de compatibilité des dimensions.

2.3 Propriétés des matrices carrées

La *trace* et le *déterminant* sont des notions intrinsèques, qui ne dépendent pas des bases de représentation choisies, mais uniquement de l'application linéaire sous-jacente.

2.3.1 Trace

Par définition, si $\mathbf{A}$ est une matrice $(p\times p)$,

$$\mathrm{tr}\mathbf{A} = \sum_{j=1}^p a_j^j,$$

et il est facile de montrer :

$$\begin{aligned}
\mathrm{tr}\alpha &= \alpha, \\
\mathrm{tr}\alpha\mathbf{A} &= \alpha\mathrm{tr}\mathbf{A}, \\
\mathrm{tr}(\mathbf{A}+\mathbf{B}) &= \mathrm{tr}\mathbf{A} + \mathrm{tr}\mathbf{B}, \\
\mathrm{tr}\mathbf{AB} &= \mathrm{tr}\mathbf{BA}, \\
& \quad \text{reste vrai si } \mathbf{A} \text{ est } (n\times p) \text{ et si } \mathbf{B} \text{ est } (p\times n) \\
\mathrm{tr}\mathbf{CC}' &= \mathrm{tr}\mathbf{C}'\mathbf{C} = \sum_{i=1}^n\sum_{j=1}^p (c_i^j)^2 \\
& \quad \text{dans ce cas, } \mathbf{C} \text{ est } (n\times p).
\end{aligned}$$

2.3.2 Déterminant

On note $|\mathbf{A}|$ le *déterminant* de la matrice carrée $\mathbf{A}$ $(p\times p)$. Il vérifie :

$$\begin{aligned}
|\mathbf{A}| &= \prod_{j=1}^p a_j^j, \text{ si } \mathbf{A} \text{ est triangulaire ou diagonale}, \\
|\alpha\mathbf{A}| &= \alpha^p|\mathbf{A}|, \\
|\mathbf{AB}| &= |\mathbf{A}||\mathbf{B}|, \\
\left|\begin{array}{cc} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{array}\right| &= |\mathbf{A}||\mathbf{C}|, \\
\left|\begin{array}{cc} \mathbf{A}_1^1 & \mathbf{A}_1^2 \\ \mathbf{A}_2^1 & \mathbf{A}_2^2 \end{array}\right| &= |\mathbf{A}_1^1||\mathbf{A}_2^2 - \mathbf{A}_2^1(\mathbf{A}_1^1)^{-1}\mathbf{A}_1^2| \qquad (1) \\
&= |\mathbf{A}_2^2||\mathbf{A}_1^1 - \mathbf{A}_1^2(\mathbf{A}_2^2)^{-1}\mathbf{A}_2^1|, \qquad (2) \\
& \quad \text{sous réserve de la régularité de } \mathbf{A}_1^1 \text{ et } \mathbf{A}_2^2.
\end{aligned}$$

Cette dernière propriété se montre en considérant les matrices :

$$\mathbf{B} = \left[\begin{array}{cc} \mathbf{I} & -\mathbf{A}_1^2(\mathbf{A}_2^2)^{-1} \\ \mathbf{0} & \mathbf{I} \end{array}\right] \text{ et } \mathbf{BAB}',$$

puis en comparant les déterminants $|\mathbf{BAB}'|$ et $|\mathbf{A}|$.

2.3.3 Inverse

L'*inverse* de $\mathbf{A}$, lorsqu'elle existe, est la matrice unique notée $\mathbf{A}^{-1}$ telle que :

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}\,;$$

elle existe si et seulement si $|\mathbf{A}| \neq 0$. Quelques propriétés :

$$(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}, \quad (\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}, \quad |\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|}.$$

2.3.4 Définitions

Une matrice carrée $\mathbf{A}$ est dite :

symétrique si $\mathbf{A}' = \mathbf{A}$,

singulière si $|\mathbf{A}| = 0$,
régulière si $|\mathbf{A}| \neq 0$,
idempotente si $\mathbf{AA} = \mathbf{A}$,
définie-positive si, $\forall \mathbf{x} \in \mathbb{R}^p, \mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$, et si $x'\mathbf{A}x = 0 \Rightarrow x = 0$,
positive, ou semi-définie-positive, si, $\forall \mathbf{x} \in \mathbb{R}^p, \mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$,
orthogonale si $\mathbf{AA}' = \mathbf{A}'\mathbf{A} = \mathbf{I}$ $(\mathbf{A}' = \mathbf{A}^{-1})$.

3 Espaces euclidiens

E est un espace vectoriel réel de dimension p isomorphe à $\mathbb{R}^p$.

3.1 Sous-espaces

- Un sous-ensemble E_q de E est un *sous-espace vectoriel* (s.e.v.) de E s'il est non vide et stable :
$$\forall(\mathbf{x}, \mathbf{y}) \in E_q^2, \forall \alpha \in \mathbb{R}, \alpha(\mathbf{x} + \mathbf{y}) \in E_q.$$
- Le q-uple $\{\mathbf{x}_1, \ldots, \mathbf{x}_q\}$ de E constitue un système *linéairement indépendant* si et seulement si :
$$\sum_{i=1}^{q} \alpha_i \mathbf{x}_i = 0 \Rightarrow \alpha_1 = \cdots = \alpha_q = 0.$$
- Un système linéairement indépendant $\mathcal{E}_q = \{\mathbf{e}_1, \ldots, \mathbf{e}_q\}$ qui engendre dans E un s.e.v. $E_q = \text{vec}\{\mathbf{e}_1, \ldots, \mathbf{e}_q\}$ en constitue une *base* et $\dim(E_q) = \text{card}(\mathcal{E}_q) = q$.

3.2 Rang d'une matrice

Dans ce sous-paragraphe, $\mathbf{A}$ est la matrice d'une application linéaire de $E = \mathbb{R}^p$ dans $F = \mathbb{R}^n$.

$$\begin{aligned}
\text{Im}(\mathbf{A}) &= \text{vect}\{\mathbf{a}^1, \ldots, \mathbf{a}^p\} \text{ est le s.e.v. de } F \textit{ image} \text{ de } \mathbf{A}\,; \\
\text{Ker}(\mathbf{A}) &= \{x \in E\,;\ \mathbf{A}x = 0\} \text{ est le s.e.v. de } E \textit{ noyau} \text{ de } \mathbf{A}\,; \\
E &= \text{Im}(\mathbf{A}) \oplus \text{Ker}(\mathbf{A}) \text{ si } \mathbf{A} \text{ est carrée associée à un endomorphisme de } E \\
\text{et } p &= \dim(\text{Im}(\mathbf{A})) + \dim(\text{Ker}(\mathbf{A})).
\end{aligned}$$

$$\begin{aligned}
\text{rang}(\mathbf{A}) &= \dim(\text{Im}(\mathbf{A})), \\
0 \leq \text{rang}(\mathbf{A}) &\leq \min(n, p), \\
\text{rang}(\mathbf{A}) &= \text{rang}(\mathbf{A}'), \\
\text{rang}(\mathbf{A} + \mathbf{B}) &\leq \text{rang}(\mathbf{A}) + \text{rang}(\mathbf{B}), \\
\text{rang}(\mathbf{AB}) &\leq \min(\text{rang}(\mathbf{A}), \text{rang}(\mathbf{B})), \\
\text{rang}(\mathbf{BAC}) &= \text{rang}(\mathbf{A}),\ \text{si } \mathbf{B} \text{ et } \mathbf{C} \text{ sont régulières}, \\
\text{rang}(\mathbf{A}) &= \text{rang}(\mathbf{AA}') = \text{rang}(\mathbf{A}'\mathbf{A}).
\end{aligned}$$

Enfin, si $\mathbf{B}$ $(p \times q)$ est de rang $q (q < p)$ et $\mathbf{A}$ est carrée $(p \times p)$ de rang p, alors la matrice $\mathbf{B}'\mathbf{AB}$ est de rang q.

3.3 Métrique euclidienne

Soit $\mathbf{M}$ une matrice carrée $(p \times p)$, symétrique, définie-positive ; $\mathbf{M}$ définit sur l'espace E :

- un *produit scalaire* : $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{M}} = \mathbf{x}'\mathbf{M}\mathbf{y}$,
- une *norme* : $\|\mathbf{x}\|_{\mathbf{M}} = \langle \mathbf{x}, \mathbf{x} \rangle_{\mathbf{M}}^{1/2}$,
- une *distance* : $\text{d}_{\mathbf{M}}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_{\mathbf{M}}$,
- des *angles* : $\cos \theta_{\mathbf{M}}(\mathbf{x}, \mathbf{y}) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{M}}}{\|\mathbf{x}\|_{\mathbf{M}} \|\mathbf{y}\|_{\mathbf{M}}}$.

La matrice $\mathbf{M}$ étant donnée, on dit que :

- une matrice $\mathbf{A}$ est $\mathbf{M}$-*symétrique* si $(\mathbf{MA})' = \mathbf{MA}$,
- deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont $\mathbf{M}$-*orthogonaux* si $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{M}} = 0$,
- un vecteur $\mathbf{x}$ est $\mathbf{M}$-*normé* si $\|\mathbf{x}\|_{\mathbf{M}} = 1$,
- une base $\mathcal{E}_q = \{\mathbf{e}_1, \ldots, \mathbf{e}_q\}$ est $\mathbf{M}$-*orthonormée* si
$$\forall (i, j), \langle \mathbf{e}_i, \mathbf{e}_j \rangle_{\mathbf{M}} = \delta_i^j.$$

3.4 Projection

Soit W un sous-espace de E et $\mathcal{B} = \{\mathbf{b}^1, \ldots, \mathbf{b}^q\}$ une base de W ; $\mathbf{P}(p \times p)$ est une matrice de projection $\mathbf{M}$-orthogonale sur W si et seulement si :
$$\forall \mathbf{y} \in E, \mathbf{Py} \in W \text{ et } \langle \mathbf{Py}, \mathbf{y} - \mathbf{Py} \rangle_{\mathbf{M}} = 0.$$

Toute matrice idempotente $(\mathbf{P}^2 = \mathbf{P})$ et $\mathbf{M}$-symétrique $(\mathbf{P}'\mathbf{M} = \mathbf{MP})$ est une matrice de projection $\mathbf{M}$-orthogonale et réciproquement.

3.4.1 Propriétés

- Les valeurs propres de $\mathbf{P}$ sont 0 ou 1 (voir § 4) :

$$\begin{array}{lll} \mathbf{u} \in W, & \mathbf{Pu} = \mathbf{u}, & \lambda = 1, \text{ de multiplicité } \dim(W), \\ \mathbf{v} \perp W, \text{(on note } \mathbf{v} \in W^{\perp}) & \mathbf{Pv} = 0, & \lambda = 0, \text{ de multiplicité } \dim(W^{\perp}). \end{array}$$

- $\mathrm{tr}\mathbf{P} = \dim(W)$.
- $\mathbf{P} = \mathbf{B}(\mathbf{B}'\mathbf{M}\mathbf{B})^{-1}\mathbf{B}'\mathbf{M}$, où $\mathbf{B} = \left[\mathbf{b}^1, \ldots, \mathbf{b}^q\right]$.
- Dans le cas particulier où les $\mathbf{b}^j$ sont $\mathbf{M}$-orthonormés :

$$\mathbf{P} = \mathbf{B}\mathbf{B}'\mathbf{M} = \sum_{i=1}^{q} \mathbf{b}^j \mathbf{b}^{j'} \mathbf{M}.$$

- Dans le cas particulier où $q = 1$ alors :

$$\mathbf{P} = \frac{\mathbf{b}\mathbf{b}'}{\mathbf{b}'\mathbf{M}\mathbf{b}}\mathbf{M} = \frac{1}{\|\mathbf{b}\|_{\mathbf{M}}}\mathbf{b}\mathbf{b}'\mathbf{M}.$$

- Si $\mathbf{P}_1, \ldots, \mathbf{P}_q$ sont des matrices de projection $\mathbf{M}$-orthogonales alors la somme $\mathbf{P}_1 + \cdots + \mathbf{P}_q$ est une matrice de projection $\mathbf{M}$-orthogonale si et seulement si : $\mathbf{P}_k\mathbf{P}_j = \delta_k^j \mathbf{P}_j$.
- La matrice $\mathbf{I} - \mathbf{P}$ est la matrice de projection $\mathbf{M}$-orthogonale sur $W^{\perp}$.

4 Eléments propres

Soit $\mathbf{A}$ une matrice carrée $(p \times p)$.

4.1 Définitions

- Par définition, un vecteur $\mathbf{v}$ définit une *direction propre* associée à une *valeur propre* λ si l'on a :

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}.$$

- Si λ est une valeur propre de $\mathbf{A}$, le noyau $\mathrm{Ker}(\mathbf{A} - \lambda\mathbf{I})$ est un s.e.v. de E, appelé sous-espace propre, dont la dimension est majoré par l'ordre de multiplicité de λ. Comme cas particulier, $\mathrm{Ker}(\mathbf{A})$ est le sous-espace propre associé, si elle existe, à la valeur propre nulle.
- Les valeurs propres d'une matrice $\mathbf{A}$ sont les racines, avec leur multiplicité, du *polynôme caractéristique* :

$$|\mathbf{A} - \lambda\mathbf{I}| = 0.$$

THÉORÈME 1. — *Soit deux matrices $\mathbf{A}(n \times p)$ et $\mathbf{B}(p \times n)$; les valeurs propres non nulles de $\mathbf{AB}$ et $\mathbf{BA}$ sont identiques avec le même degré de multiplicité. Si $\mathbf{u}$ est vecteur propre de $\mathbf{BA}$ associé à la valeur propre λ différente de zéro, alors $\mathbf{v} = \mathbf{Au}$ est vecteur propre de la matrice $\mathbf{AB}$ associé à la même valeur propre.*

Les applications statistiques envisagées dans ce cours ne s'intéressent qu'à des types particuliers de matrices.

THÉORÈME 2. — *Une matrice $\mathbf{A}$ réelle symétrique admet p valeurs propres réelles. Ses vecteurs propres peuvent être choisis pour constituer une base orthonormée de E ; $\mathbf{A}$ se décompose en :*

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}' = \sum_{k=1}^{p} \lambda_k \mathbf{v}^k \mathbf{v}^{k'}$$

où $\mathbf{V}$ est une matrice orthogonale $[\mathbf{v}^1, \ldots, \mathbf{v}^p]$ des vecteurs propres orthonormés associés aux valeurs propres λ_k, rangées par ordre décroissant dans la matrice diagonale $\boldsymbol{\Lambda}$.

THÉORÈME 3. — *Une matrice $\mathbf{A}$* réelle $\mathbf{M}$-symétrique *admet p valeurs propres réelles. Ses vecteurs propres peuvent être choisis pour constituer une base $\mathbf{M}$-orthonormée de E ; $\mathbf{A}$ se décompose en :*

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}'\mathbf{M} = \sum_{k=1}^{p} \lambda_k \mathbf{v}^k \mathbf{v}^{k'} \mathbf{M}$$

où $\mathbf{V} = [\mathbf{v}^1, \ldots, \mathbf{v}^p]$ est une matrice $\mathbf{M}$-orthogonale ($\mathbf{V}'\mathbf{M}\mathbf{V} = \mathbf{I}_p$ et $\mathbf{V}\mathbf{V}' = \mathbf{M}^{-1}$) des vecteurs propres associés aux valeurs propres λ_k, rangées par ordre décroissant dans la matrice diagonale $\boldsymbol{\Lambda}$.

Les décompositions ne sont pas uniques : pour une valeur propre simple (de multiplicité 1) le vecteur propre normé est défini à un signe près, tandis que pour une valeur propre multiple, une infinité de bases $\mathbf{M}$-orthonormées peuvent être extraites du sous-espace propre unique associé.

Le rang de $\mathbf{A}$ est aussi le rang de la matrice $\boldsymbol{\Lambda}$ associée et donc le nombre (répétées avec leurs multiplicités) de valeurs propres non nulles.

Par définition, si $\mathbf{A}$ est positive, on note la racine carrée de $\mathbf{A}$:

$$\mathbf{A}^{1/2} = \sum_{k=1}^{p} \sqrt{\lambda_k}\mathbf{v}^k\mathbf{v}^{k'}\mathbf{M} = \mathbf{V}\mathbf{\Lambda}^{1/2}\mathbf{V}'\mathbf{M}.$$

4.2 Propriétés

Si $\lambda_k \neq \lambda_j$,	$\mathbf{v}^k \perp_{\mathbf{M}} \mathbf{v}^j$;
tr$\mathbf{A} = \sum_{k=1}^{p} \lambda_k$;	$\lvert\mathbf{A}\rvert = \prod_{k=1}^{p} \lambda_k$;
si $\mathbf{A}$ est régulière,	$\forall k, \lambda_k \neq 0$;
si $\mathbf{A}$ est positive,	$\lambda_p \geq 0$;
si $\mathbf{A}$ est définie-positive,	$\lambda_p > 0$;

4.3 Décomposition en Valeurs Singulières (DVS)

Il s'agit, cette fois, de construire la décomposition d'une matrice $\mathbf{X}(n \times p)$ rectangulaire relativement à deux matrices symétriques et positives $\mathbf{D}(n \times n)$ et $\mathbf{M}(p \times p)$.

THÉORÈME 4. — *Une matrice $\mathbf{X}$ $(n \times p)$ de rang r peut s'écrire :*

$$\mathbf{X} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}' = \sum_{k=1}^{r} \sqrt{\lambda_k}\mathbf{u}^k\mathbf{v}^{k'} ; \tag{3}$$

$\mathbf{U}$ $(n \times r)$ contient les vecteurs propres $\mathbf{D}$-orthonormés ($\mathbf{U}'\mathbf{D}\mathbf{U} = \mathbf{I}_r$) de la matrice $\mathbf{D}$-symétrique positive $\mathbf{X}\mathbf{M}\mathbf{X}'\mathbf{D}$ associés aux r valeurs propres non nulles λ_k rangées par ordre décroissant dans la matrice diagonale $\mathbf{\Lambda}(r \times r)$; $\mathbf{V}$ $(p \times r)$ contient les vecteurs propres $\mathbf{M}$-orthonormés ($\mathbf{V}'\mathbf{M}\mathbf{V} = \mathbf{I}_r$) de la matrice $\mathbf{M}$-symétrique positive $\mathbf{X}'\mathbf{D}\mathbf{X}\mathbf{M}$ associés aux mêmes valeurs propres. De plus,

$$\mathbf{U} = \mathbf{X}\mathbf{M}\mathbf{V}\mathbf{\Lambda}^{-1/2} \textit{ et } \mathbf{V} = \mathbf{X}'\mathbf{D}\mathbf{U}\mathbf{\Lambda}^{-1/2}.$$

5 Optimisation

5.1 Norme d'une matrice

L'espace vectoriel E de dimension p (resp. F de dimension n) est muni de sa base canonique et d'une métrique de matrice $\mathbf{M}$ (resp. $\mathbf{D}$). Soit $\mathbf{X}$ une matrice $(n \times p)$. L'ensemble $\mathcal{M}_{n,p}$ des matrices $(n \times p)$ est un espace vectoriel de dimension np ; on le munit du *produit scalaire* :

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{\mathbf{M},\mathbf{D}} = \mathrm{tr}\mathbf{X}\mathbf{M}\mathbf{Y}'\mathbf{D}. \tag{4}$$

Dans le cas particulier où $\mathbf{M} = \mathbf{I}_p$ et $\mathbf{D} = \mathbf{I}_n$, et en notant $\mathrm{vec}(\mathbf{X}) = \left[\mathbf{x}^{1'}, \ldots, \mathbf{x}^{p'}\right]'$ la matrice “vectorisée”, ce produit scalaire devient :

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{\mathbf{I_p},\mathbf{I_n}} = \mathrm{tr}\mathbf{X}\mathbf{Y}' = \sum_{i=1}^{n}\sum_{j=1}^{p} x_i^j y_i^j = \mathrm{vec}(\mathbf{X})'\mathrm{vec}(\mathbf{Y}).$$

La *norme* associée à ce produit scalaire (4) est appelée norme trace :

$$\begin{aligned} \|\mathbf{X}\|^2_{\mathbf{M},\mathbf{D}} &= \mathrm{tr}\mathbf{X}\mathbf{M}\mathbf{X}'\mathbf{D}, \\ \|\mathbf{X}\|^2_{\mathbf{I_p},\mathbf{I_n}} &= \mathrm{tr}\mathbf{X}\mathbf{X}' = \mathrm{SSQ}(\mathbf{X}) = \sum_{i=1}^{n}\sum_{j=1}^{p}(x_i^j)^2 \\ &\quad \text{(SSQ signifie “sum of squares”).} \end{aligned}$$

La *distance* associée à cette norme devient, dans le cas où $\mathbf{D}$ est une matrice diagonale ($\mathbf{D} = \mathrm{diag}(w_1, \ldots, w_n)$), le critère usuel des *moindres carrés* :

$$\mathrm{d}^2(\mathbf{X}, \mathbf{Y}) = \|\mathbf{X} - \mathbf{Y}\|^2_{\mathbf{M},\mathbf{D}} = \sum_{i=1}^{n} w_i \|\mathbf{x}_i - \mathbf{y}_i\|^2_{\mathbf{M}}.$$

5.2 Approximation d'une matrice

Les matrices $\mathbf{X}$, $\mathbf{M}$ et $\mathbf{D}$ sont définies comme ci-dessus ; $\mathbf{X}$ est supposée de rang r. On cherche la matrice $\mathbf{Z}_q$, de rang q inférieur à r, qui soit la plus proche possible de $\mathbf{X}$.

THÉORÈME 5. — *La solution du problème :*

$$\min_{\mathbf{Z}} \left\{ \|\mathbf{X} - \mathbf{Z}\|^2_{\mathbf{M},\mathbf{D}} ; \mathbf{Z} \in \mathcal{M}_{n,p},\ \textit{rang}(\mathbf{Z}) = q < r \right\} \tag{5}$$

est donnée par la somme des q premiers termes de la décomposition en valeurs singulières (3) de $\mathbf{X}$:

$$\mathbf{Z}_q = \sum_{k=1}^{q} \sqrt{\lambda_k}\mathbf{u}^k\mathbf{v}^{k'} = \mathbf{U}_q\mathbf{\Lambda}_q^{1/2}\mathbf{V}'_q.$$

Le minimum atteint est :

$$\|\mathbf{X} - \mathbf{Z}_q\|^2_{\mathbf{M},\mathbf{D}} = \sum_{k=q+1}^{r} \lambda_k.$$

Les matrices $\mathbf{U}_q, \mathbf{\Lambda}_q$ et $\mathbf{V}_q$ contiennent les q premiers vecteurs et valeurs propres donnés par la DVS de $\mathbf{X}$; $\mathbf{Z}_q$ est appelée approximation de rang q de $\mathbf{X}$.

Ce théorème peut se reformuler d'une manière équivalente. On note $\widehat{\mathbf{P}_q}$ (resp. $\widehat{\mathbf{Q}_q}$) la projection $\mathbf{M}$-orthogonale sur $E_q = \operatorname{Im}(\mathbf{V}_q)$ (resp. $\mathbf{D}$-orthogonale sur $F_q = \operatorname{Im}(\mathbf{U}_q)$) :

$$\begin{aligned}
\widehat{\mathbf{P}_q} &= \sum_{k=1}^{q} \mathbf{v}^k\mathbf{v}^{k'}\mathbf{M} = \mathbf{V}_q\mathbf{V}'_q\mathbf{M} \\
\widehat{\mathbf{Q}_q} &= \sum_{k=1}^{q} \mathbf{u}^k\mathbf{u}^{k'}\mathbf{D} = \mathbf{U}_q\mathbf{U}'_q\mathbf{D}, \\
\mathbf{Z}_q &= \widehat{\mathbf{Q}_q}\mathbf{X} = \mathbf{X}\widehat{\mathbf{P}_q}'.
\end{aligned}$$

PROPOSITION 6. — *Avec les notations précédentes :*

$$\begin{aligned}
\widehat{\mathbf{P}_q} &= \textit{arg}\ \max_{\mathbf{P}_q} \left\{ \|\mathbf{X}\mathbf{P}'_q\|^2_{\mathbf{M},\mathbf{D}} ; \right. \\
&\qquad \left. \mathbf{P}_q \textit{ projection } \mathbf{M}\textit{-orthogonale de rang } q < r \right\}, \\
\widehat{\mathbf{Q}_q} &= \textit{arg}\ \max_{\mathbf{Q}_q} \left\{ \|\mathbf{Q}_q\mathbf{X}\|^2_{\mathbf{M},\mathbf{D}} ; \right. \\
&\qquad \left. \mathbf{Q}_q \textit{ projection } \mathbf{D}\textit{-orthogonale de rang } q < r \right\}.
\end{aligned}$$