Analyse en composantes principales

Présenté par : M. HAMMAD

l'ACP est l'une des plus anciennes méthodes factorielles. Elle a été conçue par Karl Peason (1901) et intégrée à la statistique par Harold Hotelling (1933). Elle est utilisée lorsqu'on observe n individus sur p variables quantitatives $X_1, X_2, \cdots, X_p$, $(p \leq n)$ présentant des liaisons multiples que l'on veut analyser. Ces observations sont regroupées dans un tableau (matrice) rectangulaire X ayant p lignes (variables) et n colonnes (individus) :

$$
\begin{array}{cc}
 & \begin{array}{cccccc} X_1 & X_2 & \cdots & \cdots & X_n & \\ \downarrow & \downarrow & & & \downarrow & \end{array} \\
X = & \left(\begin{array}{cccccc} x_{11} & x_{12} & \cdots & \cdots & x_{1n} \\ x_{21} & \ddots & & & x_{2n} \\ \vdots & & & x_{ij} & \vdots \\ \vdots & & & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & \cdots & x_{pn} \end{array}\right) \begin{array}{c} \leftarrow \; X^1 \\ \leftarrow \; X^2 \\ \vdots \\ \vdots \\ \leftarrow \; X^p \end{array}
\end{array}
$$

0.1 Principe de l'ACP

Rappelons que l'objectif principal est d'obtenir une représentation fidèle du nuage des individus de $\mathbb{R}^p$ en le projetant sur un espace de faible dimension. Le choix de l'espace de projection s'effectue selon le critère de l'inertie, i.e. on cherche le sous-espace de dimension k portant l'inertie maximale du nuage.
But : Classer un ensemble I d'individus, $\{$individus "proches", individus "éloignés"$\}$.
On suppose que I est muni d'une distance d_I.

Définition 0.1.1.

Soit (I, d_I) *un espace métrique, et on note* (η, E, M) *le triplé formé par :*
η *: nuage de points de* E
E *: espace euclidien*
M *: une métrique sur* E*,*
on dit que (η, E, M) *est une image euclidienne de* (I, d_I) *s'il existe :*

$$\begin{array}{ll} h : (I, d_I) & (E, M) \\ \quad (i, i^{'}) & h(i, i^{'}) = d_I(i, i^{'}) = \|h(i) - h(i^{'})\|_M \end{array}$$

Définition 0.1.2.

On dit que deux images euclidienne (η_1, E_1, M_1) *et* (η_2, E_2, M_2) *sont équivalentes, s'il existe une bijection* $f : \eta_1 \longrightarrow \eta_2$; *tel que* $\|x_i - x_{i'}\|_{M_1} = \|f(x_i) - f(x_{i'})\|_{M_2}$.

Remarque 0.1.1.

Parmis toutes les images euclidiennes équivalents de (I, d_I)*, la plus simple serait celle pour la quelle la* $\dim(E)$ *est la plus faible.*

Pour l'ACP : étant donnée une image euclidienne (η, E, M), il s'agit alors de la "simplifier", i.e., trouver une image euclidienne equivalente (η_1, E_1, M_1) tel que : $\dim E_1 < \dim E$.

Exemple 0.1.1.

Soit I *l'ensemble des individus d'écrits par troix variables (la taille, le poid et l'âge) on peut observer notre nuage* I *graphiquement. Pour ce nuage en peut fair un ACP c'est à dire de changer l'image de nuage vers une autre plus claire et on peut réduire la dimension à une autre plus faible à condition de garder les mêmes propriétés.*

Remarque 0.1.2.

On générale, impossible d'observer les nuages sur des espaces de dimension supérieure à 3, c'est pour celà on utilise l'ACP.

Les démarches pour simplifier le nuage (η, E, M)

Pour simplifier le nuage (η, E, M), on se posera les questions suivantes :

1. Quel est le point de E le plus "proche" de η?.
2. Quelle est la droite de E la plus "proche" de η?.
3. Quel est le plan affine de E le plus proche de η?.
4. $\vdots$
5. Quel est le sous espace affine de E de dimension k le plus "proche" de η?.

Hypothèse 0.1.1.

On suppose que l'espace des individu E est muni d'une métrique M sur I ensemble de n individus d'écrit par un ensemble J de p variables quauntitatives.

Chaque individu i est muni d'un poid : $p_i > 0$; $\sum_{i=0}^{n} P_i = 1$.

L'espace des variables F est muni d'une métrique N tel que $N = D_p$ (voir (3.1))

Alors, on obtient un tableau de donnée X avec le schéma de dualité suivant :

$$\left\{\begin{array}{l} V = XD_pX', \\ W = X'MX, \end{array}\right. \quad D_p(f_i, f_{i'}) = p_i\delta_{ii'} \;\; et \;\; \delta_{ii'} = \left\{\begin{array}{ll} 1 & i = i' \\ 0 & i \neq i' \end{array}\right.$$

Remarque 0.1.3.

Nécessité de définir un indice (un critère) qui permet de mesurer la proximité :

- *D'un point de E et η.*
- *D'une droite de E et η.*
- *D'un sous-espace affine de E et η.*

0.2 Inertie et moment d'inertie

0.2.1 Inertie de η par rapport à un point

Définition 0.2.1.

L'inertie de η par rapport à un point a est $I_a = \sum_{i=1}^{n} p_i \|x_i - a\|_M^2$.

Elle mesure l'approximité de η et a, de plus on a, $I_a = I_g + \|g - a\|$.

Donc, si $a = g, I_a = I_g$ est minimale.

Remarque 0.2.1.

g est le point de E le plus "proche" de η au sens de l'inertie : $I_g = \sum_{i=1}^{n} p_i \|x_i\|_M^2$.

I_g : mesure la dispersion autour du centre de gravité, g et I_g généralisent les notions de moyenne et de variance.

Autre expression de I_g

On a :

$$\begin{array}{rcl} \|x_i\|_M^2 & = & M(x_i, x_i) \\ & = & x_i' M x_i \text{ (scalaire)}, \end{array}$$

d'où :

$$\begin{array}{rcl} tr(x_i' M x_i) & = & x_i' M x_i \\ & = & x_i x_i' M, \end{array}$$

l'inertie totale

$$\begin{aligned} I_g &= \sum_{i=1}^{n} p_i \|x_i\|_M^2 \\ &= tr(\sum_{i=1}^{n} p_i x_i x_i') \\ &= tr(VM), \quad VM: \text{ est un endomorphisme de}\, E^*. \end{aligned}$$

0.2.2 Moment d'inertie par rapport à une droite

Soit Δu une droite engendré par un vecteur $u \in E$ normé ($\|u\|_M = 1$), soit $\Delta u^{\perp}$ le supplimentaire M-orthogonale de Δu dans E, $E = \Delta u \oplus \Delta u^{\perp}$. On a $x_i \in E \Longrightarrow x_i = \alpha_i + \beta_i$,

$$\begin{aligned} I_{\Delta u} &= \sum_{i=1}^{n} p_i \|x_i - \alpha_i\|_M^2 \\ &= \sum_{i=1}^{n} p_i \|\beta_i\|_M^2. \\ I_{\Delta u^{\perp}} &= \sum_{i=1}^{n} p_i \|x_i - \beta_i\|_M^2 \\ &= \sum_{i=1}^{n} p_i \|\alpha_i\|_M^2. \end{aligned}$$

On sait que $\alpha_i \in \Delta u$ est la projection sur Δu qui s'écrit : $\alpha_i = c_i u$ tel que $c_i \in \mathbb{R}$ est la coordonnée de projection de x_i sur la droite Δu :

$$c_i = M(x_i, u),$$

on définit la forme lineaire suivante :

$$\begin{aligned} v : E &\longrightarrow \mathbb{R} \\ x &\longrightarrow v(x) = M(x, u) \end{aligned}$$

on note $v = Mu \in E^*$, autrement dit :

$$\begin{aligned} I_{\Delta u^{\perp}} = \sum_{i=1}^{n} p_i \|\alpha_i\|_M^2 &= \sum_{i=1}^{n} p_i M(\alpha_i, \alpha_i) \\ &= \sum_{i=1}^{n} p_i M(c_i u, c_i u) \\ &= \sum_{i=1}^{n} p_i c_i^2 M(u, u), \end{aligned}$$

or $M(u,u) = \|u\|_M^2 = 1$, donc $I_{\Delta u^\perp} = \sum_{i=1}^{n} p_i c_i^2$.

Soit C le vacteur qui contient tous les coordonnée de projection de x_i sur la droite Δ_u,

$$C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} M(x_1,u) \\ M(x_2,u) \\ \vdots \\ M(x_n,u) \end{pmatrix} = \begin{pmatrix} x_1Mu. \\ x_2Mu \\ \vdots \\ x_nMu \end{pmatrix} \Longrightarrow C = X'Mu.$$

Expression matricielle de l'inertie $I_{\Delta u^\perp}$

Proposition 0.2.1.

$$I_{\Delta u^\perp} = MVM(u,u).$$

Preuve 0.2.1.

$$\begin{aligned} I_{\Delta u^\perp} = \sum_{i=1}^{n} p_i c_i^2 &= \|C\|_{D_p}^2 \\ &= D_p(C,C) \\ &= C'D_pC \\ &= (X'v)'D_p(X'v) \quad \textit{car } C = X'Mu \textit{ et } v = Mu \Longrightarrow C = X'v \\ &= vXD_pX'v \\ &= XD_pX'(v,v) \\ &= V(v,v) \quad \textit{car } V = XD_pX' \\ &= V(Mu,Mu) \ \textit{car } v = Mu \\ &= MVM(u,u), \end{aligned}$$

$\Longrightarrow I_{\Delta u^\perp} = MVM(u,u)$.

Remarque 0.2.2.

On a, $I_{\Delta u} = I_g - I_{\Delta u^\perp}$, *par conséqunce on obtient :* $I_{\Delta u} = Tr(VM) - MVM(u,u)$.

Remarque 0.2.3.

1. *Si* $\|u\| \neq 1 \Longrightarrow I_{\Delta u^\perp} = \dfrac{MVM(u,u)}{M(u,u)}$.

2. *Si* $$\begin{aligned} \eta \subset I_{\Delta u^\perp} \Longrightarrow I_{\Delta u^\perp} = 0 &\Longrightarrow MVM(u,u) = 0 \quad \Longrightarrow M(u,VMu) = 0 \\ &\Longrightarrow VMu = 0 \\ &\Longrightarrow u \in Ker(VM). \end{aligned}$$

La construction du sous-espace vectoriel de $\dim k$ le plus "proche" de η

Soit W est un sous-espace vectoriel de E tel que $W = W_1 \oplus W_2$ M-ourthogonal à $W^{\perp}$ alors,

$$W^{\perp} = W_1^{\perp} \oplus W_2^{\perp} \quad et$$

$I_g = I_W + I_{W^{\perp}} = I_W + I_{W_1^{\perp}} + I_{W_2^{\perp}}$.

Tout sous-espace vectoriel de dimension "$k+1$", W_{k+1} tel que $I_{W_{k+1}}$ est minimum contient un espace vecoriel W_k de dimension "k" et tel que I_{W_k} soit minimum.

$$W_{k+1} = W_k + \Delta u,$$

Δu : *est une droite de E (un espace vectoriel de dimension 1).*

Preuve 0.2.2.

Soit F un sous-espace de dimension k. Comme $\dim(W_{k+1}^{\perp}) + \dim(F) = p-1, W_{k+1}^{\perp}$ *et F ont au moins une direction commune.*

Soit $\Delta u_{k+1} \subset W_{k+1}^{\perp} \cap F, (u_{k+1} \neq 0)$, On peut alors écrire $\widetilde{F} = F \oplus \Delta u_{k+1}$, où F est le supplémentaire $M-$orthogonal de Δu_{k+1} dans $\widetilde{F}$.

$\widetilde{F}$ est de dimension $k+1$, et par définition de W_{k+1} on a donc $I_{\tilde{F}} \geq I_{W_{k+1}}$. Par ailleurs, par définition,on a aussi $I_{\Delta u_{k+1}} \geq I_{\Delta u}$. Ainsi,

$$I_F = I_{\tilde{F}} - I_{\Delta u_{k+1}} \geq I_{W_{k+1}} - I_{\Delta u} = I_{W_k}$$

Remarque 0.2.4.

l'idée du théoreme (0.2.2), nous permet de déduie qu'à partir d'un sous espace W_k de dimension "k". (tel que I_{W_k} est minimum) de construire un sous espace vectoriel W_{k+1} de dimension "$k+1$" de $I_{W_{k+1}}$ minimale est l'augmentant d'une droite.

Remarque 0.2.5.

On peut déduire que, si l'espace $W_k^{\perp}$ est d'inertie maximale alors, $W_{k+1}^{\perp}$ est aussi reste d'inertie maximale,

$$W_{k+1}^{\perp} = W_k^{\perp} \oplus \Delta u_{k+1}.$$

Preuve 0.2.3.

Soit F un sous-espace de dimension $k+1$. Comme $\dim(W_k^{\perp}) + \dim(F) = p+1, W_k^{\perp}$ *et F ont au moins une direction commune. Soit $\Delta u \subset W_k^{\perp} \cap F, (u \neq 0)$. On peut alors écrire $F = \widetilde{F} \oplus \Delta u$, où $\widetilde{F}$ est le supplémentaire M-orthogonal de Δu dans F. $\widetilde{F}$ est de dimension k, et par déffnition de $W_k^{\perp}$ on a donc $I_{\widetilde{F}} \leq I_{W_k^{\perp}}$. Par ailleurs, par définition de u_{k+1}, on a aussi $I_{\Delta u} \leq I_{\Delta u_{k+1}}$. Ainsi,*

$$I_F = I_{\tilde{F}} + I_{\Delta u} \leq I_{W_k^{\perp}} + I_{\Delta u_{k+1}} = I_{W_{k+1}^{\perp}}.$$

Calcul des axes principaux

Commençons par trouver l'axe principal Δu *si* u *est* M*-normé on a vu que* $I_{\Delta u^\perp} = MVM(u,u) = u'MVMu$, *si non, en notant* $v = \dfrac{u}{\|u\|_M}$,
$I_{\Delta v^\perp} = I_{\Delta u^\perp} = v'MVMv = \dfrac{u'MVMu}{\|u\|_M^2} = \dfrac{u'MVMu}{u'Mu}$.

On veut maximiser cette quantité en $u \in \mathbb{R}^p$. *Pour cela, on commence par chercher les points critiques de la fonction considérée, i.e., les points où le gradient est nul :*

$$\nabla_u \left(\frac{u'MVMu}{u'Mu} \right) = \frac{2MVMu}{u'Mu} - \frac{(u'MVMu)(2Mu)}{(u'Mu)^2}$$

u est donc solution de

$$MVMu = \frac{(u'MVMu)\,(Mu)}{(u'Mu)} = I_{\Delta u^\perp} Mu \Longleftrightarrow VMu = I_{\Delta u^\perp} u,$$

puisque M *est inversible. Autrement dit,* u *est vecteur propre de la matrice* VM *(* VM *est endomorphisme voir (remarque 3.2.2)) associé à la valeur propre* $I_{\Delta u^\perp}$
Pour maximiser $I_{\Delta u^\perp}$ *il faut donc choisir pour le vecteur propre* u *associé à la plus grande valeur propre* λ *de la matrice* VM. *On a alors* $I_{\Delta u^\perp} = \lambda$.

Remarque 0.2.6.
Ce résultat se généralise aux autres axes principaux.

Remarque 0.2.7.

$$\begin{aligned} I_{\Delta u^\perp} &= MVM(u,u) \\ &= M(u, VMu) \\ &= M(u, \lambda u) \\ &= \lambda M(u,u) \\ &= \lambda. \end{aligned}$$

u est le vecteur propre associé à la plus grande valeur propre de VM *on ce cas la on dit que le minimum est atteint.*

Conclusion 0.2.1.
si λ *est la plus grande valeur propre alors la majorété des individus sont très proches de la droite* Δu.

Propriété 0.2.1.

1. VM *est* M *symétrique* $\Longrightarrow (MVM)' = MVM$
2. *tout les valeurs propres de* MV *sont réells et positives.*
3. *si* k *est le nombre de valeurs propres distanctes alors* $E = E_1 \oplus E_2 \oplus \cdots \oplus E_k$ *où* E_i *est le sous espace propre associé à* λ_i.
   Si $i \neq j E_i \perp_M E_j$, $E_i \cap E_j = \{0\}$.

Conclusion 0.2.2.

On peut construire une nouvelle base $\{u_i;\, i = 1, \cdots, i = p\}$ *de* E *formée uniquement de vecteurs propres* u_i *de* VM.

0.3 Axes et plans principaux d'une ACP

premier axe principal

Soit $\{u_i/i = 1, \cdots, p\}$ *la base de* E *formée de vecteurs propres de* VM, *soit* $u \in E$, $u = \sum_{i=1}^{p} \alpha_i u_i,\ \ \alpha_i \in \mathbb{R}$.

Définition 0.3.1.

Le 1^{er} *axe principal est l'axe* Δu *engendré par le vecteur normé* u, *tel que* $I_{\Delta u}$ *soit minimale ou bien* $I_{\Delta u^{\perp}}$ *est maximal, i.e.*

$$\begin{cases} I_{\Delta u} \textit{ est maximale} \\ \textit{avec } M(u,u) = 1. \end{cases}$$

$$\begin{aligned} M(u,u) &= M\left(\sum_{i=1}^{p} \alpha_i u_i, \sum_{j=1}^{p} \alpha_j u_j\right), \ \textit{car } u = \sum_{i=1}^{p} \alpha_i u_i \textit{ tel que } \alpha_i \in \mathbb{R} \\ &= \sum_{j.i}^{p} \alpha_i \alpha_j M\left(u_i, u_j\right) \\ &= \sum_{i=1}^{p} \alpha_i^2 M\left(u_i, u_i\right) \ \textit{ car si } i \neq j, M\left(u_i, u_j\right) = 0 \\ &= \sum_{i=1}^{p} \alpha_i^2 = 1. \end{aligned}$$

Calcul de $I_{\Delta u^\perp}$

$$\begin{aligned}
I_{\Delta u^\perp} &= MVM(u,u)\\
&= M(u,VMu)\\
&= M\left(\sum_{i=1}^{p}\alpha_i u_i, \sum_{j=1}^{p}\alpha_j\lambda_j u_j\right),\ \ \textit{car}\ VMu = VM\left(\sum_{j=1}^{p}\alpha_j u_j\right) = \left(\sum_{j=1}^{p}\alpha_j VMu_j\right) = \sum_{j=1}^{p}\alpha_j\lambda_j u_j\\
&= \sum_{i,j}^{p}\alpha_i\alpha_j\lambda_j M\left(u_i,u_j\right)\\
&= \sum_{i=1}^{p}\alpha_i^2\lambda_i M\left(u_i,u_i\right)\ \ \textit{car si}\ i\neq j, M\left(u_i,u_j\right)=0\\
&= \sum_{i=1}^{p}\alpha_i^2\lambda_i.
\end{aligned}$$

Remarque 0.3.1.

Pour $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$*, on veut maximiser* $I_{\Delta u_1^\perp}$ *sous la containte* $\sum_{i=1}^{p}\alpha_i^2 = 1$*, c'est-à-dire maximiser* $\sum_{i=1}^{p}\lambda_i\alpha_i^2$ *alors :*

$$\sum_{i=1}^{p}\lambda_i\alpha_i^2 \leq \sum_{i=1}^{p}\lambda_1\alpha_i^2 = \lambda_1\sum_{i=1}^{p}\alpha_i^2 = \lambda_1.$$

La solution est :

$$\alpha_1 = 1, \alpha_2 = \alpha_3 = \cdots = \alpha_p = 0.$$

Le 1^{er} *axe pricipal coïncide avec l'axe engendré par le vecteur propre associé à la plus grande valeur propre* λ_1.

Au premier axe principal Δu_1 *on associé :*
Dans E^* *:*
"le premier facteur" principal $v_1 = Mu_1$

$$\begin{aligned}
v_1 : E &\longrightarrow R\\
x_i &\longrightarrow v_1\left(x_i\right) = M\left(x_i,u_1\right)
\end{aligned}$$

$M\left(x_i,u_1\right) = C_{1i}$ *: coordonnée de projection de* x_i *sur* Δu_1 *(*i^{eme} *individu sur* 1^{er} *axe)*

Dans *F :*
$C^1 = X'v_1$

$C^1 = X'Mu_1 = \begin{pmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1n} \end{pmatrix}$

C'est la 1^{ere} *composante principale (est une variable qui contient les coordonnées de "n" individus).*

$$\begin{aligned} Var(C^1) &= \|C^1\|^2_{Dp} = \lambda_1 \\ &= MVM(u_1, u_1) \\ &= I_{\Delta u_1^{\perp}}. \end{aligned}$$

Part d'inertie expliquée par Δu_1

$I_{\Delta u_1} = Tr(VM) - \lambda_1$ *avec* $I_{\Delta u_1^{\perp}} = MVM(u_1, u_1) = \lambda_1,$

la Part d'inertie expliquée par Δu_1 *est le rapport entre l'inertie du sous espace supplimentair et l'inertie totale :* $\dfrac{\lambda_1}{Tr(VM)}$.

cas particulier : *Si* $I_{\Delta u_1} = 0$, *i.e.,* $tr(VM) = \sum_{i=1}^{p} \lambda_i = \lambda_1 \Longrightarrow \eta \subset \Delta u_1$ *et* $\dfrac{\lambda_1}{Tr(VM)} = 1.$

0.3.1 Plan principal

Définition 0.3.2.
C'est le sous-espace vectoriel de dimension 2 et de moment d'inertie miminum, il est engendré par :
-Un premier axe principal Δu_1
-Un deuxiéme axe Δu *M-ortogonal à* Δu_1 *et tel que* $I_{\Delta u}$ *soit minimale (*$I_{\Delta u_1^{\perp}}$ *soit maximale).*

Checher u tel que :

-$\begin{cases} \Delta u \perp \Delta u_1 \\ et \\ I_{\Delta u_1^{\perp}} \text{ est Maximale} \end{cases}$

-On a

$M(u, u) = 1 \Longleftrightarrow \sum_{i=1}^{p} \alpha_i^2 = 1$ *et*

$$\begin{aligned} M(u, u_1) &= M(\sum_{i=1}^{p} \alpha_i u_i, u_1) \\ &= \sum_{i=1}^{p} \alpha_i M(u_i, u_1) \\ &= \alpha_1. \end{aligned}$$

$M(u_i, u_1) = 0$ *pour* $i \neq 1, M(u, u_1) \ = 0 \Longleftrightarrow \alpha_1 = 0.$

Maximiser $I_{\Delta u_1^{\perp}}$ *sous la contrainte* $\sum_{i=2}^{p} \alpha_i^2 = 1,$

C'es-à-dire maximiser $\sum_{i=2}^{p} \lambda_i \alpha_i^2$ *sous* $\sum_{i=2}^{p} \alpha_i^2 = 1$ *comme :* $\lambda_i \leq \lambda_2 \ \forall i = \overline{2, p}$:

$$\sum_{i=2}^{p} \lambda_i \alpha_i^2 \leq \sum_{i=2}^{p} \lambda_i \alpha_i^2 = \lambda_2 \sum_{i=2}^{p} \alpha_i^2 = \lambda_2.$$

La solution est :

Le maximum est λ_2, *il est atteint si* $\alpha_2 = 1$, *et* $\alpha_3 = \cdots = \alpha_p = 0.$

Le 2^{eme} *axe pricipal coincïde avec l'axe engendré par le vecteur propre associé 'a la deuxième plus grande valeur propre* λ_2 *de* VM.

On associé à Δu_2 :

- ***Dans*** E^* *: "le premier facteur" principle* $v_2 : Mu_2$

$$\begin{aligned} v_2 : E &\longrightarrow \mathbb{R} \\ x_i &\longrightarrow v_2(x_i) = M(x_i, u_2) \end{aligned}$$

$M(x_i, u_2) = C_{2i}$: *la coordonnée de projection de* x_i *sur* Δu_2 *(* i^{eme} *individu sur* 2^{eme} *axe)*

-***Dans*** F *:*

$C^2 = X'v_2$

$$C^2 = X'Mu_2 = \begin{pmatrix} c_2 1 \\ c_2 2 \\ \vdots \\ c_2 n \end{pmatrix}$$

C'est le 2^{eme} *composante principale (est une variable qui contient les coordonnée de "n" individu).*

$$\begin{aligned} Var(C^2) &= \|C^2\|_{Dp}^2 = \lambda_2 \\ &= MVM(u_2, u_2) \\ &= I_{\Delta u_2}^1. \end{aligned}$$

Part d'inertie du plans principal

On note $\mathcal{P}$ *le plans principal tel que :* $\mathcal{P} = \Delta u_1 \oplus \Delta u_2$

$$\begin{aligned} I_{\mathcal{P}} &= I_q - I_{\mathcal{P}\perp} \\ &= I_g - \left(I_{\Delta u_1^{\perp}} + I_{\Delta u_1^{\perp}} \right) \\ &= Tr(VM) - (\lambda_1 + \lambda_2). \end{aligned}$$

La part d'inertie expliqué par le plans principal est :

$$\frac{\lambda_1 + \lambda_2}{Tr(VM)}.$$

Généralisation

Le sous-espace principal de $\dim k$

Définition 0.3.3.

Le sous-espace principal de dimension k est engendré par les k premiers axes principaux : $\Delta u_1, \Delta u_2, \cdots, \Delta u_k$ engendrés par les vecteurs propres associées aux k plus grandes valeurs propres $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ de VM.
Au j^{eme} axe principl Δu_j est associe :

- ***Dans*** E^* *: le j^{eme} facteur principal $v_j = Mu_j$ tel que $C_{ji} = M(x_i, u_j) = \langle x_i, u_j \rangle_M$.*

- ***Dans F*** *:la j^{eme} composante principal $c^j = X'v_j = X'Mu_j = \begin{pmatrix} c_{j1} \\ c_{j2} \\ \vdots \\ c_{jn} \end{pmatrix}$*

$I_{\Delta u_j^{\perp}} = MVM(u_j, u_j) = \lambda_j$.

La part d'inertie expliquée par le sous espace principal de dimension k est :

$$\frac{\sum_{j=1}^{k} \lambda_j}{\sum_{j=1}^{p} \lambda_j} = \sum_{j=1}^{k} \frac{\lambda_j}{\mathrm{Tr}(VM)} = \sum_{j=1}^{k} \mathrm{part}\left(\Delta u_j\right).$$

Remarque 0.3.2.

$E_k = \Delta u_1 \oplus \Delta u_2 \oplus \cdots \oplus \Delta u_k$
$I_{E_k} = I_g - I_{E_k^{\perp}} = \mathrm{Tr}(VM) - (\lambda_1 + \lambda_2 + \cdots + \lambda_k)$, *i.e.*,

$$I_{E_k} = 0 \Longrightarrow \eta \subset E_k \Longrightarrow \frac{\sum_{j=1}^{k} .\lambda_j}{\mathrm{Tr}(VM)} = 1.$$

0.3.2 Répresentation graphique des individus

On doit calculer les coordonnées de projection des individus sur le plan principal, on doit donc connaître les cooronnées de projertion de x_j sur Δu_1 et Δu_2 qui sont contenues dans C^1 et C^2

$$C^1 = X'Mu_1 = \begin{pmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1i} \\ \vdots \\ c_{1n} \end{pmatrix} \begin{matrix} \longleftarrow (c_{11}, c_{21}) \longrightarrow \\ \longleftarrow (c_{12}, c_{22}) \longrightarrow \\ \\ \longleftarrow (c_{1i}, c_{2i}) \longrightarrow \\ \\ \longleftarrow (c_{1n}, c_{2n}) \longrightarrow \end{matrix} \begin{pmatrix} c_{21} \\ c_{22} \\ \vdots \\ c_{2i} \\ \vdots \\ c_{2n} \end{pmatrix} = X'Mu_2 = C^2$$

tel que $\|u_1\| = \|u_2\| = 1$

0.3.3 Répresentation des variables

Définition 0.3.4.

On s'intéresse dans cette section au nuage $\mathcal{V}$ des variables centrées qui sont des éléments de l'espace $\mathbb{R}^n$. Pour obtenir de bonnes représentations plane de ce nuage de points, on peut adopter la méme démarche que pour le nuage des individus, et faire une ACP dans l'espace des variables. L'objectif est de trouver les sous-espaces principaux $F_1, F_2, \cdots$, etc de $\mathbb{R}^n$ qui conservent au mieux l'information liée à l'inertie contenue dans le nuage des variables $\mathcal{V}$. On calcul :

1. *Le tableau transposé X' de X.*
2. *la Matrice des poids D_p*
3. *La Matrice de variance covarianse $W = X'MX$.*
4. *les valeurs propres de WD_p.*
   On a : $VMu_i = XD_pX'Mu_i = \lambda_i u_i$. En multipliant cette identité par $X'M$, et en utilisant l'identité $c^i = X'Mu_i$.
   On obtient :
   $$X'MXD_rX'Mu_i = X'M\lambda_i u_i \Longleftrightarrow WD_pc^i = \lambda_j c^i,$$
   alors les composante principal c^i sont les vecteurs propres de WD_p associée à $\lambda_1, \lambda_2, \cdots, \lambda_p$
5. *WD_p est un opérateur lineaire de F.*
6. *soit $\left\{ \frac{c^i}{\|c^i\|_{D_p}} / i = 1, \cdots, p \right\}$ une base D_p-orthonormée de vecteur propres de WD_p qui sont egales $\left\{ \frac{c^i}{\sqrt{\lambda_i}} / i = 1, \cdots, p \right\}$ car $\|c^i\|_{D_p} = \lambda_i$ sachants que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$*

(a) *on définit* $d_i = \dfrac{c^i}{\sqrt{\lambda_i}}$ *et* Δd_i *le* i^{eme} *axe principle tel que*

$$I_{\Delta d_i^{\perp}} = \|c_i\|_M^2 = \sum_{j=1}^{p} c_{ji}^2 = M\left(c_i, c_i\right) = D_p W D_p \left(d_i, d_i\right) = \lambda_i$$

(b) *le facteur principal* $w_i = D_p d_i$ *tel que*

$$C_{ji} = D_p\left(x^j, d_i\right),$$

(c) *les composantes principals*

$$c_i = X w_i = X D_p d_i = D_p\left(x^j, \frac{c^i}{\sqrt{\lambda_i}}\right) = \frac{D_p\left(x^j, c^i\right)}{\|c^i\|_{D_p}}$$

(d) *part d'inertie expliquée par le sous espace* F_k

$$\sum_{j=1}^{k} \frac{\lambda_j}{\operatorname{Tr}\left(W D_p\right)} = \frac{\sum_{j=1}^{k} \lambda_j}{\operatorname{Tr}\left(W D_p\right)}$$

Cas particulier représentation des variables dans $\mathbb{R}^2$

On doit calculer les coordonnée C_1, C_2 *de projection des variables* $x^1, x^2, \cdots, x^n$ *sur le plan principal* $\Delta c^1, \Delta c^2$ *par :*

$$C_1 = \frac{D_p\left(x^j, C^1\right)}{\|C^1\|_{D_p}}, \quad C_2 = \frac{D_p\left(x^j, C^2\right)}{\|C^2\|_{D_p}}.$$

0.3.4 Représentation conjointe

Il s'agit de représenter les individus sur Δu_1 *et* Δu_2 *pour répresenter les variables sur même plans, on doit calculer les coordonnées de projection* M*-orthogonal des variables sur* Δu_1 *et* Δu_2 *par :*

$$\beta_{j1} = M(e_j, u_1)\ , \beta_{j2} = M(e_j, u_2).$$

La qualitée de représentation des individus

Dans une carte des individus, on ne peut tirer de conclusions sur les individus (regroupement, individus exceptionnels, etc...) que si ces individus sont bien représentés dans le plan principal considéré.

Définition 0.3.5.
La qualité de représentation de l'individu i sur le plans principal est mesurée par le cosinus carré de l'angle que fait x_i avec sa projection α_i sur $\mathcal{P}$ tel que $\mathcal{P} = \Delta u_1 \oplus \Delta u_2$:

$$\cos^2(\alpha) = \frac{\|\alpha_i\|_M}{\|x_i\|_M},$$

on sait que $\alpha_i = c_{1i}u_1 + c_{2i}u_2$ car $\alpha_i \in \mathcal{P}$ alors
$\|\alpha_i\|_M^2 = \|c_{1i}u_1\|_M^2 + \|c_{2i}u_2\|_M^2 = \|c_{1i}\|_M^2 + \|c_{2i}\|_M^2, \|c_{ji}\|_M^2 = c_{ji}^2$ *car* c_{ji} *est costante,*
et par suit :

$$\cos^2(\alpha) = \frac{c_{1i}^2 + c2i^2}{\|x_i\|_M^2}.$$

Remarque 0.3.3.
- *Si* $\cos^2(\alpha)$ *est proche de 1 il est donc bien représenté sur le plans car il est un peu proche de* $\mathcal{P}$.
- *Si* $\cos^2(\alpha)$ *est proche de 0 il est donc mai représenté sur le plans car il est un peu proche de* $\mathcal{P}^{\perp}$.

Contribution des points au facteur(CPF)

On a :

$$\begin{aligned} I_{\Delta}u_j^{\perp} = \lambda_j = \|C^j\|_{D_p}^2 &= D_p(C^j, C^j) \\ &= \sum_{i=1}^{n} p_i c_{ji}^2, \end{aligned}$$

alors le CPF est $\dfrac{p_i c_{ji}^2}{\lambda_j} \leq 1$,
On peut conclure que c'est le CPF est faible donc les points sont trés proche de l'origine si non ils sontéloignées.

Contribution des points à l'inertie totale(INR)

$$INR = \frac{p_i\|x_i\|_M^2}{I_g} \leq 1,$$

comme le CPF si l'INR est faible les points sont trés proches de l'origine si nonils sont éloigniées.

Tableau récapitulatif

	Individus	*Variables*
Espace vectoriel	$\mathbb{R}^p$	$\mathbb{R}^n$
Tableau des données	$X'(n,p)$	$X(p,n)$
Matrice des poids	$D_p = diag(p_1, p_2, \cdots, p_n)$	M
Métrique	M	D_p
Matrice à diagonaliser	$VM = XD_pX'M$	$WD_p = X'MXD_p$
Valeurs propres non nulles	$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$	$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$
Axes principaux	$u_1, u_2, \cdots, u_p$ $\begin{cases} XD_pX'Mu_j = \lambda_j u_j \\ < u_j, u_i >_M = \delta_{ji} \end{cases}$	$d_1, d_2, \cdots, d_p$ $\begin{cases} X'MXD_pd_i = \lambda_i d_i \\ < d_j, d_i >_{D_p} = \delta_{ji} \end{cases}$
Composantes principales	$c^j = X'Mu_j$ $\|c^j\|_D^2 = Var(c^j) = \lambda_j$	$c_i = XD_pd_i$
Facteurs principaux	$d_1, d_2, \cdots, d_p; d_j = \dfrac{c^j}{\sqrt{\lambda_j}}$	$u_1, u_2, \cdots, u_p$

0.4 Cas particulier : ACP normée, Cercle des corrélations

Exemple préliminaire :

On considère n individu sur les quels on a mesurè deux variable x^1 *et* x^2 *telque :*
x^1 *: poids en Kg*
x^2 *: taille en M,*
la matrice des variance covariance est donnée par :
$V = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}$, *si* $M = I_2$, $\quad \Delta e_1, \Delta e_2$ *telque* $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

$$\begin{cases} I_{\Delta e_1^\perp} = V(e_1, e_2) = 4 \\ I_{\Delta e_2^\perp} = V(e_2, e_3) = 3 \end{cases} \Longrightarrow I_g = Tr(VM) = 7.$$

Part d'inerté expliquée par Δe_1 *est :* $\frac{4}{7}$.

Part d'inerté expliquée par Δe_2 *est :* $\frac{3}{7}$.

Remarque 0.4.1.
Les deux axes sont d'égale importance.

Exprimons maintenant, le poids en Hg, la matrice V *est :*

$$V = \begin{pmatrix} 400 & 20 \\ 20 & 3 \end{pmatrix}.$$

$$\begin{cases} I_{\Delta e_1^{\perp}} = V(e_1, e_2) = 400 \\ I_{\Delta e_2^{\perp}} = V(e_2, e_3) = 3 \end{cases} \quad I_g = tr(VM) = 403.$$

et par suit :

Part d'inerté expliquée par Δe_1 *est :* $\dfrac{400}{403}$.

Part d'inerté expliquée par Δe_2 *est :* $\dfrac{3}{403}$.

Remarque 0.4.2.

Le deuxième axe peut être négligeable.

Conclusion 0.4.1.

Si $M = I_p$ *les résultat dépendent des unités de mesure des variables.*

On remarque que les unités de mesures ont une influence sur la représentation graphique d'ACP dans ce cas là on vas entamer l'ACP normée.

Le choix de la métrique M *est toujour délicat :seul l'utilisateur peut définir correctement la notion de distance entre individus.*

Prendre $M = I_p$ *revient à travailler sur la matrice* V *des variances-covariaances il n'y a pas alors de distinction, entre axes principaux et facteurs principaux. Cependant, les résultats obtenus ne sont pas invariants si l'on change linéairement l'unité de mesure des variables. Les covariances sont multipliées par un facteur* K*, la variance par un facteur* K^2 *si l'on choisit une unité de mesure* K *fois plus petite pour une variable.*

Le choix de $M = D_{1/\sigma^2}$ *est le plus communément fait et a pour conséquence de rendre les distances entre individus invariantes par transformation linéaire séparée de chaque variable et de s'affranchir des unités de mesure ce qui est particuliérement intéressant lorsque les variables sont hétérogénes.*

On sait que l'usage de cette métrique à réduit les variables par l'écart-type.

En partique on travaillera donc sur le tableau centré-réduit Z *associé à* X *et on utilisera la métrique* $M = I_p$.

Comme la matrice des données est centrée et réduite est la matrice de corrélation R*, les facteur principaux seront donc les vecteurs propres successifs de* R *rangés selon l'ordre décroissant des valeur propres.*$Ru = \lambda u$ *avec* $\|u\|^2 = 1$.

La première composante principale C*(et les autres sous la contrainte d'orthogonalité) est la combinaison linéaire des variables centrées er réduites ayant une variance maximale* $C = Z'u$.

Apercu sur le schéma de dualité

Remarque 0.4.3.

Faire l'ACP de $(X, D_{1/\sigma^2}, V)$ *est équivalence à faire l'ACP de* (Z, I_p, R) *l'ACP normée avec* $M = D_{1/\sigma^2}, Z = D_{1/\sigma}.X$

$$\|z_i - z_i'\|^2_{I_p} = \|x_i - x_i'\|^2_{D_{1/\sigma^2}} I_p(z_i, z_i') = D_{1/\sigma^2}(x_i, x_i').$$

Preuve 0.4.1.

$$\begin{aligned} D_{1/\sigma^2}(x_i, x_i') &= x_i'.D_{1/\sigma^2}.x_i' \\ &= x_i'.D_{1/\sigma}.Id_p.D_{1/\sigma}.x_i' \\ &= z_i'.Id_p.z_i' = Id_p(z_i, z_i'). \end{aligned}$$

Remarque 0.4.4.

On sait que :

$$\begin{aligned} VMu_j = \lambda_j u_j &\Longrightarrow VD_{1/\sigma^2}u_j = \lambda_j u_j \, , M = D_{1/\sigma^2}. \\ &\Longrightarrow VD_{1/\sigma}D_{1/\sigma}.u_j = \lambda_j u_j. \\ &\Longrightarrow D_{1/\sigma}VD_{1/\sigma}D_{1/\sigma}.u_j = \lambda_j D_{1/\sigma}u_j. \\ &\Longrightarrow RD_{1/\sigma}.u_j = \lambda_j D_{1/\sigma}u_j \ \ , R = D_{1/\sigma}VD_{1/\sigma}. \end{aligned}$$

Donc $\bigsqcup = D_{1/\sigma}.u_j$ *est le vecteur propre de* R *associé à* λ_j.

Conclusion 0.4.2.

Si $M = D_{1/\sigma}$ *les résultat ne dépendent pas des unités de mesuré des variables.*

Représentation des variables

On sait que :

$$C_k = \frac{D_p(x^j, C^k)}{\|C^k\|_{D_p}} = \frac{D_p(x^j, C^k)}{\sqrt{\lambda_k}} = r(x^j, c^k) - 1 \leq \frac{D_p(x^j, C^k)}{\sqrt{\lambda_k}} \leq 1,$$

toutes les variables sont située sur la sphére unité δ_n *de l'espce des variables* $\mathbb{R}^n$. *L'intersection de cette sphére avec le premier plan factoriel est donc un cercle unité appelé cercle des corrélations.*

Remarque 0.4.5.

$$\begin{aligned} r(x^j, C^k) &= \frac{D_p(x^j, C^k)}{\sqrt{\lambda_k}} \\ &= \frac{[XD_pC^k]_j}{\sqrt{\lambda_k}} \\ &= \frac{[XD_pX'Mu_k]_j}{\sqrt{\lambda_k}} \\ &= \frac{[VMu_k]_j}{\sqrt{\lambda_k}} \\ &= \frac{[\lambda_k u_k]_j}{\sqrt{\lambda_k}}, \\ &= \sqrt{\lambda_k}[u_k]_j, \end{aligned}$$

$\sqrt{\lambda_k}[u_k]_j$: *contient les cofficient de corrélation linéaire de* x^j *avec la* k^{eme} *composante*

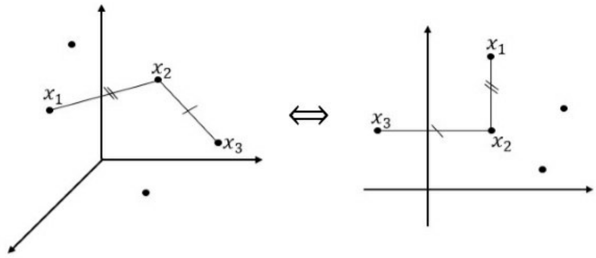

FIGURE 1 – Image euclidienne-équivalente.

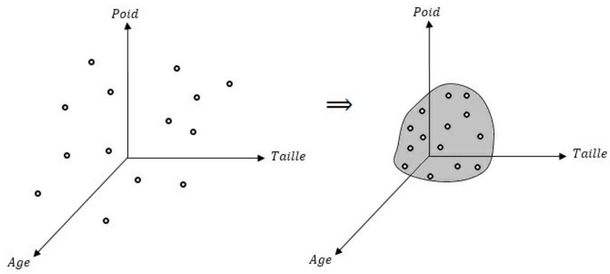

FIGURE 2 – ACP-Dimension faible.

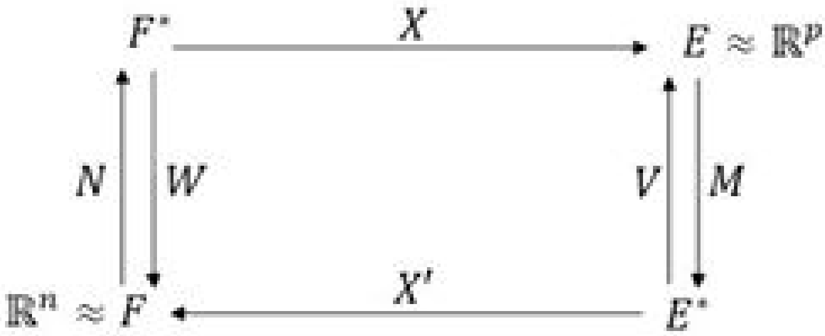

FIGURE 3 – Schéma de dualité.

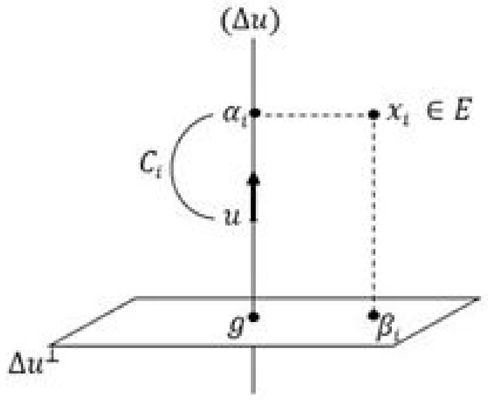

FIGURE 4 – Proximitée par rapport à une droite.

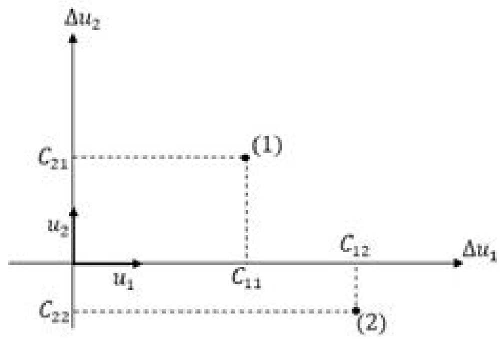

FIGURE 5 – Représentation des individeus.

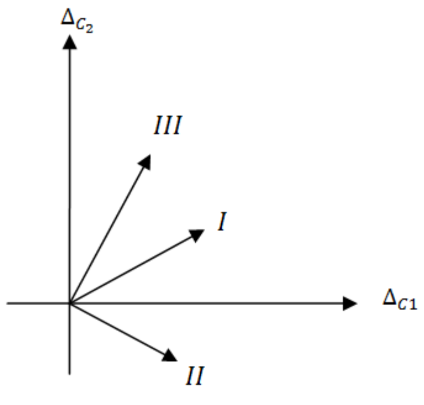

FIGURE 6 – Repésentation des variables.

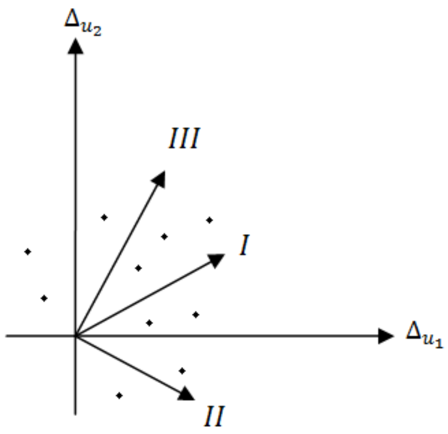

FIGURE 7 – Représentation conjointe.

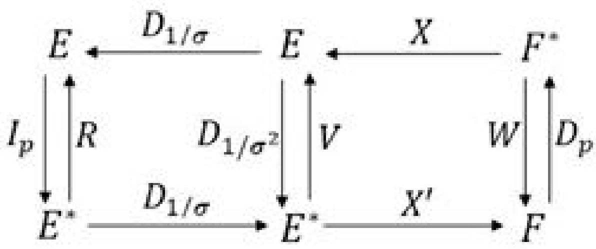

FIGURE 8 – Schéma de dualité ACP normé.

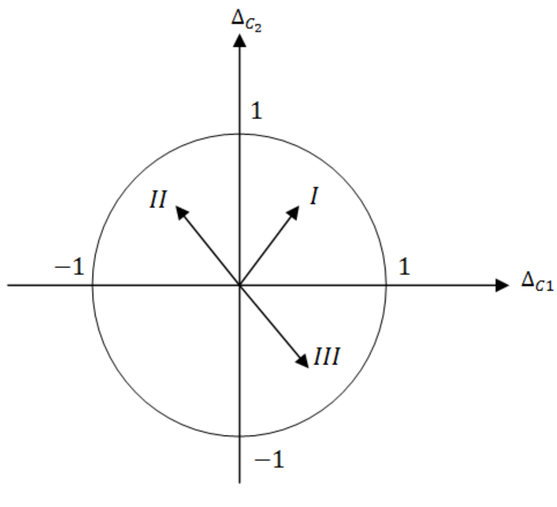

FIGURE 9 – Cercle de corrélation.