ACP

Déf: le phénomène étudié apparait sous forme de données numériques. C'est possible regrouper ces données dans un tableau et interpréter le tableau comme une matrice:

Par exemple, considérons la matrice X de type (p,q):

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1q} \\ x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2q} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{iq} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pj} & \cdots & x_{pq} \end{pmatrix}$$

C'est possible de décomposer cette matrice en p lignes L_i $(i = 1, \dots, p)$ et q colonnes C_j $(j = 1, \dots, q)$:

$$X = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_p \end{pmatrix} = (C_1, C_2, \dots, C_j, \dots, C_q)$$

La transposée de X notée X' sera alors:

$$X' = (L_1, L_2, \dots, L_i, \dots, L_p) = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_q \end{pmatrix}$$

Souvent on interprete :

- la colonne C_j comme les p observations de la variable j
- la ligne L_i comme les valeurs des q variables pour la $i^{\text{ème}}$ observation :

$$L_i = (x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{iq})$$

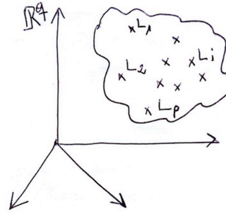
On peut ainsi représenter :

- la colonne C_j par un point dans l'espace de dimension p ($\mathbb{R}^p$) ;
- la ligne L_i par un point dans l'espace de dimension q ($\mathbb{R}^q$).

La matrice X peut être interprétée alors comme donnée par :

- q points C_j en $\mathbb{R}^p$ (c'est-à-dire que chaque point admet p composantes) ou
- p points L_i en $\mathbb{R}^q$ (c'est-à-dire que chaque point admet q composantes).

Par la suite, nous nous situons dans cette dernière situation.



L'analyse des données de composantes principales consiste en l'étude des projections des points du nuage sur un axe, un plan, ou un hyperplan judicieusement déterminé. Mathématiquement, l'analyse en composantes principales serait le meilleur ajustement du nuage par un sous-espace vectoriel en $\mathbb{R}^q$.

Ajustement du nuage :

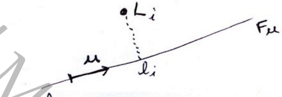
1) coordonnées d'un point en $\mathbb{R}^q$:

Soit un point L_i de $\mathbb{R}^q$:

$$L_i = (x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{iq})$$

Soit un axe F_u engendré par un vecteur-colonne quelconque u, de norme 1 ; c'est-à-dire :

$$\|u\| = \sqrt{u.u} = \sqrt{\sum_{i=1}^{n} u_i^2} = u'.u = 1$$



l_i désigne la coordonnée de L_i sur F_u, c'est-à-dire la projection de Li sur F_u.

Par définition, la coordonnée l_i est égale au produit scalaire des vecteurs L_i et u :

$$l_i = L_i \cdot u = \sum_{j=1}^{q} x_{ij} u_j$$

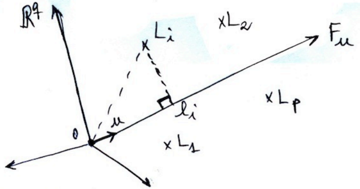
selon

2) Ajustement du nuage par un axe suivant la méthode des moindres carrés :

Soit le nuage des p points $L_1, L_2, \dots, L_i, \dots, L_p$ en $\mathbb{R}^q$.

Soit l'axe F_u, engendré par le vecteur unitaire u passant par l'origine des coordonnées :

$$u'.u = 1$$



a) Nous projetons les points du nuage sur l'axe F_u, de façon que la somme des carrées des distances des points à l'axe soit minimale, c'est-à-dire que :

$$\sum_{i=1}^{p} \overline{L_i l_i}^2 \text{ soit minimale}$$

Alors, minimiser $L_i l_i$ équivalent à minimiser $\overline{Ol_i}^2$ dans le triangle rectangle Ol_iL_i.

On sait que $\quad \overline{Ol_i} = L_i \cdot u$

donc $\quad \overline{Ol_i}^2 = (L_i \cdot u)'(L_i \cdot u)$

ou $\quad \overline{Ol_i}^2 = u' L'_i L_i u$

et $\quad \sum_{i=1}^{p} \overline{Ol_i}^2 = u'\left(\sum_{i=1}^{p} L'_i L_i\right) u$

Mais $\quad \sum_{i=1}^{p} L'_i L_i = X' X$

Finalement, il s'agit de minimiser $\quad u'X'Xu$ sous la condition $\quad u' \cdot u = 1$

b) Utilisons la méthode du multiplicateur de Lagrange :

Soit Φ la ~~forme~~ fonction de Lagrange dans laquelle λ_1 est le multiplicateur de Lagrange.

$$\Phi = u'X'Xu - \lambda_1(u'u - 1) \; ;$$

dérivons Φ par rapport à u : $\quad \dfrac{\partial \Phi}{\partial u} = 2X'Xu - 2\lambda_1 u$

égalisons cette dérivée à zéro ; on obtient :

$$(X'X)u = \lambda_1 u$$

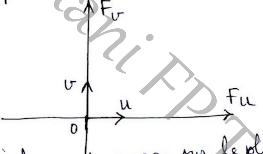
c'est_à_dire que u est un vecteur propre de la matrice $X'X$

En plus $\quad u'X'Xu = u'\lambda_1 u = \lambda_1 u'u = \lambda_1 \cdot 1 = \lambda_1$.

Ainsi le maximum de $u'X'Xu$, sous la condition $u'u = 1$, correspond au valeur propre λ_1 maximale de la matrice $X'X$.

3) Ajustement du nuage par un plan selon la méthode des moindres carrés :

Maintenant, nous nous proposons d'ajuster le nuage par un plan, déterminé par l'axe F_u précédent et par un axe F_v, de vecteur unitaire v (c-à-d $v'v = 1$) passant par l'origine des coordonnées et perpendiculaire à l'axe F_u (c-à-d que $v'u = 0$)



La meilleur projection du nuage sur le plan ainsi définie, équivalente cette fois-ci à maximiser $v'X'Xv$ sous la condition $v'v = 1$ et aussi $v'u = 0$

Soit Ψ la fonction forme de Lagrange dans laquelle λ_2 et μ_2 sont les multiplicateurs de Lagrange :

$$\Psi = v'X'Xv - \lambda_2(v'v - 1) - \mu_2(v'u - 0)$$

Dérivons Ψ par rapport à v :

$$\frac{\partial \Psi}{\partial v} = 2X'Xv - 2\lambda_2 v - \mu_2 u$$