

$$X = \begin{pmatrix} 3 & -1 \\ -1 & 2 \\ 4 & 2 \end{pmatrix} \quad X' = \begin{pmatrix} 1 & -2 \\ -3 & 1 \\ 2 & 1 \end{pmatrix} \quad V = \begin{pmatrix} 14/3 & -1 \\ -1 & 2 \end{pmatrix} \quad X'' = \begin{pmatrix} \sqrt{3/14} & -\sqrt{2} \\ -3\sqrt{3/14} & \sqrt{2}/2 \\ 2\sqrt{3/14} & \sqrt{2}/2 \end{pmatrix} \quad R = \begin{pmatrix} 1 & -\frac{\sqrt{3}}{2\sqrt{7}} \\ -\frac{\sqrt{3}}{2\sqrt{7}} & 1 \end{pmatrix}$$

1. Composantes Principales – Solution

Un tableau des vecteurs propres de V est donné ci-contre.

v_1	v_2
-3	-1
1	-3

- 1-1) Faut-il réaliser une ACP normée ?
- 1-2) Déduisez les valeurs propres de V .
- 1-3) La 1ère composante principale permet-elle de bien représenter le nuage X ?
- 1-4) Calculez le tableau des vecteurs unitaires U définissant les composantes principales.
- HW) Vous vérifierez (par produit matriciel) qu'ils sont bien unitaires.
- HW) Vous vérifierez qu'ils sont V -orthogonaux (par produit matriciel).
- 1-5) Projetez x'_2 sur u_1 , puis généralisez à tous les individus de sorte de définir C'_1 .
- 1-6) Généralisez à toutes les composantes, c'est-à-dire projetez X' sur U , afin d'obtenir le tableau C' .
- HW) Vous calculerez (avec la formule pratique en dim. 1) la variance de chaque composante principale.
- 1-7) Soit $D = I/n$. Calculez ${}^t C' D C'$.
- 1-8) Par quel tableau faut-il diviser C' (élément par élément) pour le réduire ?
- 1-9) Déduisez le tableau C'' des composantes principales normées.
- HW) Vous calculerez ${}^t C'' D C''$. Était-il utile de faire ce calcul ?

2. Composantes Principales – Analyse des Variables

- 2-1) Tracez un cercle de rayon 1, centré à l'origine.
- 2-2) Projetez X''_1 le long de C''_2 .
- 2-3) Généralisez (par produit matriciel) de sorte de projeter le nuage X'' dans l'espace des composantes normées.
- 2-4) Qu'obtient-on en fait (puisque les variables et les composantes sont centrées-réduites) ?
- 2-5) Placez les variables réduites dans le cercle des corrélations (plan 1-2).
- 2-6) Sont-elles au bord du cercle ?
- 2-7) Quantifiez la qualité de représentation de X_2 par C'_1 , puis C'_2 , et enfin $\{C'_1, C'_2\}$. Commentez cette dernière.
- 2-8) Sur quelle(s) variable(s) s'appuiera-t-on pour expliquer chaque composante ?

3. Composantes Principales – Analyse des Individus

- 3-1) Tracez un repère de base $\{u_1, u_2\}$ et placez les individus projetés.
- 3-2) Si vous ne l'avez pas fait après la séquence TD1-TP1, vous devez avoir calculé le produit matriciel $X'^t X'$ avant le TD2 afin d'obtenir (sur la diagonale) les (carrés des) normes des individus centrés. Pourquoi aurait-on pu aussi les obtenir à l'aide de $C'^t C'$?
- HW) Vous calculerez $C'^t C'$ pour vérifier.
- 3-3) Quantifiez la qualité de représentation de x_2 sur C'_1 . Aurait-on pu se prévoir ce résultat ?
- 3-4) Pouvez-vous proposer une construction (matricielle et/ou tableau) permettant de généraliser à tous les individus et toutes les composantes ? Faites le calcul.
- 3-5) Sur quel(s) individus(s) s'appuiera-t-on pour expliquer chaque composante ?

L'Analyse en Composantes Principales sur ces données n'a aucun d'intérêt pour des raisons assez évidentes :

- on ne sait pas ce que sont les variables, ni pour quels individus on les a observées ; il est alors impossible de donner du sens aux composantes
- n est petit et surtout $p = 2$; à quoi bon réduire une si petite dimensionnalité ?

Si on se place du seul point de vue technique, on peut néanmoins se prêter à l'exercice consistant à :

- (HW) choisir combien de composantes retenir 1-3)
- (HW) pour chacune d'entre elle, choisir quel(s) individus et quelle(s) variables utiliser pour l'interpréter 2-8) et 3-5)