Corrigé du devoir surveillé d'Analyse des données (mars 2018)

Exercice 1

1)
$$g = \frac{1}{5}(15\ 40\ 65\ 90)' = (3\ 8\ 13\ 18\ 23)' = x_{i_3}.$$

$$Y = \begin{pmatrix} \dots \\ x_i - g \\ \dots \end{pmatrix} = \begin{pmatrix} -2 & \dots & -2 \\ -1 & \dots & -1 \\ 0 & \dots & 0 \\ 1 & \dots & 1 \\ 2 & \dots & 2 \end{pmatrix}$$

2) Toutes les lignes de Y étant égales, le rang de Y est égal à 1, et par conséquent celui de V_Y est aussi 1. Il en résulte qu'il existe un et un seul axe factoriel non trivial. Donc toutes les données sont situées sur cet unique axe factoriel.

rang égal à 1, donc un seul axe factoriel non trivial.

3) On
$$V_Y \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 10 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$
, c'est donc un vecteur directeur d'un axe factoriel. Etant

donné qu'il est de norme égale à
$$5: u = \frac{1}{\sqrt{5}} \begin{pmatrix} 1\\1\\1\\1 \end{pmatrix}$$
 est le premier et unique axe factoriel associé à la valeur propre $\lambda = 10$.

4) L'inertie totale du nuage étudié $I_T = tr(V_Y) = 10$.

5) La première composante principale de l'individu
$$i_5$$
:

$$\Psi_1^{i_5} = (y_{i_5})'u = \frac{1}{\sqrt{5}}(2+2+2+2+2) = \frac{10}{\sqrt{5}} = 4.472$$

6) Il n'existe qu'un seul "axe factoriel" pour représenter les variables. Les coordonnées des variables sur cet axe sont celles du vecteur $\eta = \sqrt{\lambda}u$. Donc, sur cet unique axe, chaque variable a la même coordonnée qui vaut $\frac{\sqrt{10}}{\sqrt{5}} = 1.414$.

1