Correction Exo 2

On considère la matrice des données

$$X = \begin{pmatrix} 4 & 3 & 6 \\ 2 & 5 & 8 \\ 0 & 1 & 7 \end{pmatrix}, \qquad n = 3,\ p = 3.$$

1. $$\bar{X}^{(1)} = \frac{4+2+0}{3} = \frac{6}{3} = 2, \qquad \bar{X}^{(2)} = \frac{3+5+1}{3} = \frac{9}{3} = 3, \qquad \bar{X}^{(3)} = \frac{6+8+7}{3} = \frac{21}{3} = 7.$$

   Le vecteur des moyennes (transposé) est donc

   $$\bar{x}^t = (2,\ 3,\ 7).$$

   La matrice des moyennes $\bar{X}$ (dimension 3×3) a chaque ligne égale à $\bar{x}^t$.

2. La matrice centrée est $X - \bar{X}$; ses colonnes sont les vecteurs centrés.

   $$X - \bar{X} = \begin{pmatrix} 4-2 & 3-3 & 6-7 \\ 2-2 & 5-3 & 8-7 \\ 0-2 & 1-3 & 7-7 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 1 \\ -2 & -2 & 0 \end{pmatrix}.$$

   Notons ces colonnes :

   $$(X-\bar{X})^{(1)} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}, \quad (X-\bar{X})^{(2)} = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix}, \quad (X-\bar{X})^{(3)} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

3. On utilise la définition (normalisation par n donnée dans l'énoncé) :

   $$V(X) = \frac{1}{n}(X-\bar{X})^\top (X-\bar{X}).$$

   Calcul des produits scalaires colonne à colonne (ou par multiplication matricielle) donne

   $$V(X) = \begin{pmatrix} \frac{8}{3} & \frac{4}{3} & -\frac{2}{3} \\ \frac{4}{3} & \frac{8}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix} \approx \begin{pmatrix} 2.6667 & 1.3333 & -0.6667 \\ 1.3333 & 2.6667 & 0.6667 \\ -0.6667 & 0.6667 & 0.6667 \end{pmatrix}.$$

   Les variances (diagonale) sont donc

   $$\mathrm{Var}(X^{(1)}) = \frac{8}{3}, \qquad \mathrm{Var}(X^{(2)}) = \frac{8}{3}, \qquad \mathrm{Var}(X^{(3)}) = \frac{2}{3}.$$

4. Le théorème donne

   $$\mathrm{Cov}(X^{(1)}, X^{(3)}) = \frac{1}{n}\langle X^{(1)}, X^{(3)}\rangle - \bar{X}^{(1)}\bar{X}^{(3)}.$$

   Ici $\langle X^{(1)}, X^{(3)}\rangle = 4\cdot 6 + 2\cdot 8 + 0\cdot 7 = 24 + 16 + 0 = 40$. Donc

   $$\mathrm{Cov}(X^{(1)}, X^{(3)}) = \frac{40}{3} - 2\cdot 7 = \frac{40}{3} - 14 = \frac{40-42}{3} = -\frac{2}{3}.$$

   (Concorde avec l'élément $(1,3)$ de $V(X)$ ci-dessus.)

5. La corrélation de Bravais–Pearson entre les variables i, j est

   $$r(X^{(i)}, X^{(j)}) = \frac{\mathrm{Cov}(X^{(i)}, X^{(j)})}{\sigma_i \sigma_j}, \quad \sigma_k = \sqrt{\mathrm{Var}(X^{(k)})}.$$

On obtient (ici les variances sont simples fractions) :

$$\sigma_1 = \sigma_2 = \sqrt{\tfrac{8}{3}}, \qquad \sigma_3 = \sqrt{\tfrac{2}{3}}.$$

Les coefficients de corrélation sont donc

$$R(X) = \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}.$$

(On peut vérifier numériquement que, par exemple, $r(X^{(1)}, X^{(3)}) = (-2/3)\big/\big(\sqrt{8/3}\sqrt{2/3}\big) = -1/2$.)

6. Les vecteurs colonnes centrés sont les vecteurs de $\mathbb{R}^3$ donnés ci-dessus. La corrélation entre deux variables est le cosinus de l'angle entre leurs vecteurs centrés :

$$r(X^{(i)}, X^{(j)}) = \cos\angle\big((X - \bar{X})^{(i)}, (X - \bar{X})^{(j)}\big).$$

Ainsi :
— $r(X^{(1)}, X^{(2)}) = \frac{1}{2}$ correspond à un angle $\theta_{12} = \arccos(1/2) = 60°$;
— $r(X^{(1)}, X^{(3)}) = -\frac{1}{2}$ correspond à un angle $\theta_{13} = \arccos(-1/2) = 120°$;
— $r(X^{(2)}, X^{(3)}) = \frac{1}{2}$ correspond à un angle $\theta_{23} = 60°$.