TD5 : Analyse en composantes principales

Exercice 1

On observe $p = 2$ variables sur $n = 5$ individus uniformément pondérés.

x^1	x^2
3	-2
3	2
0	-1
4	-3
0	-1

1. Calculez la matrice S de covariance.
2. Diagonalisez la matrice S.
3. Calculez la matrice C des composantes principales.

Exercice 2

On observe $p = 2$ variables sur $n = 4$ individus de même poids.

	x	y
a	2	2
b	-2	-2
c	-1	1
d	1	-1

1. Représentez les individus a, b, c et d dans le plan.
2. Calculez la matrice de covariance S associée aux observations en utilisant la formule matricielle.
3. Diagonalisez la matrice S.
4. Faites apparaître les vecteurs propres sur le graphique de la question 1.
5. Calculez la matrice C des composantes principales et faites le lien avec le graphique.

Exercice 3

Au mois de novembre 2012, les services météorologiques ont relevé les précipitations p (en cm), la température maximale t_{max} (en ˚C) et la température minimale t_{min} (en ˚C) dans différentes villes de France. Voici les résultats obtenus pour les six villes uniformément pondérées que nous considérerons :

	p	t_{max}	t_{min}
Ajaccio	12.04	23.7	5.9
Brest	17.18	15.5	-1.8
Dunkerque	11.83	13.1	2.8
Nancy	6.23	13.5	-2.4
Nice	16.99	21.1	7.2
Toulouse	3.87	20.3	-0.9

1. Calculez les moyennes $\overline{p}$, $\overline{t}_{max}$ et $\overline{t}_{min}$. Donnez la matrice X des données centrées.
2. Calculez la matrice de covariance Σ.
3. Un logiciel de statistique nous donne les matrices D et P telles que $\Sigma = PD\,{}^tP$,

$$D = \begin{bmatrix} 32.71 & 0 & 0 \\ 0 & 18.57 & 0 \\ 0 & 0 & 3.94 \end{bmatrix} \qquad \text{et} \qquad P = \begin{bmatrix} 0.73 & 0.62 & 0.29 \\ 0.40 & -0.73 & 0.56 \\ 0.56 & -0.29 & -0.78 \end{bmatrix}.$$

   Quelles sont les valeurs propres de Σ ? Expliquez pourquoi la matrice P est inversible et $P^{-1} = {}^tP$.
4. Calculez la matrice des composantes principales C et représentez les villes dans le plan principal.
5. Calculez les corrélations linéaires entre les variables initiales et les deux premières composantes principales. Représentez les résultats sur le cercle des corrélations.
6. Interprétez la position des villes dans le plan principal et commentez.