Chap II - Parametres d'un nuage de points.

exemple.

prix de 3 produits x, y, z dans 3 magasins differents, 1, 2, 3.

	x	y	z
magasin 1	3	1,2	5
magasin 2	4	2	6
magasin 3	3	1,8	6

x : lait
y : eau
z : lait.

I- Parametres statistiques.

A : matrice des données

$$A = \begin{pmatrix} 3 & 1,2 & 5 \\ 4 & 2 & 6 \\ 3 & 1,8 & 6 \end{pmatrix}$$

n : nombre d'individus.

Parametres	Definition	Exemple
Moyenne	$\overline{X} = E(X) = \frac{1}{n} \Sigma x$	$\overline{X} = \frac{10}{3}$.
Variance	$V(X) = E[(X-\overline{x})^2] = \frac{1}{n} \Sigma (x - \overline{x})^2$	$V(X) = \frac{2}{9}$
Covariance	$Cov(X,Y) = \Sigma((X-\overline{X})(Y-\overline{Y})) = \frac{1}{n} \Sigma (x-\overline{x})(y-\overline{y})$	$Cov(X,Y) = \frac{1}{15}$
Correlation	$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_x \times \sigma_y}$.	

→ variables centrées
$$X^c = X - \overline{X} \Rightarrow \overline{X}^c = 0.$$
→ variables centrées et reduites (unités différentes)
$$X^s = \frac{X - \overline{X}}{\sigma_x} \Rightarrow \overline{X}^s = 0$$
$$V(X^s) = 1.$$

II- Approche geometrique.

a. matrice de variance - covariance.

Soit A : matrice des variables

A^c : matrices des variables centrées

On appelle matrice de variance - covariance, la matrice $\Sigma = \frac{1}{n} {}^t A^c . A^c$.

$$A = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} \qquad A^c = \begin{pmatrix} x_1 - \overline{X} & y_1 - \overline{Y} & z_1 - \overline{Z} \\ x_2 - \overline{X} & y_2 - \overline{Y} & z_2 - \overline{Z} \\ x_3 - \overline{X} & y_3 - \overline{Y} & z_3 - \overline{Z} \end{pmatrix}$$

$$\Sigma = \frac{1}{3} \begin{pmatrix} x_1 - \overline{X} & x_2 - \overline{X} & x_3 - \overline{X} \\ y_1 - \overline{Y} & y_2 - \overline{Y} & y_3 - \overline{Y} \\ z_1 - \overline{Z} & z_2 - \overline{Z} & z_3 - \overline{Z} \end{pmatrix} \begin{pmatrix} x_1 - \overline{X} & y_1 - \overline{Y} & z_1 - \overline{Z} \\ x_2 - \overline{X} & y_2 - \overline{Y} & z_2 - \overline{Z} \\ x_3 - \overline{X} & y_3 - \overline{Y} & z_3 - \overline{Z} \end{pmatrix}$$

$$\Sigma = \frac{1}{3} \begin{pmatrix} (x_1 - \overline{X})^2 + (x_2 - \overline{X})^2 + (x_3 - \overline{X})^2 \\ (y_1 - \overline{Y})(x_1 - \overline{X}) + (y_2 - \overline{Y})(x_2 - \overline{X}) + (y_3 - \overline{Y})(x_3 - \overline{X}) \\ \end{pmatrix}$$

$$\boxed{\Sigma = \begin{pmatrix} V(X) & cov(X,Y) & cov(X,Z) \\ cov(Y,X) & V(Y) & cov(Y,Z) \\ cov(Z,X) & cov(Z,Y) & V(Z) \end{pmatrix}}$$

Remarque : * Σ est une matrice carrée et symetrique
* $tr(\Sigma)$: somme des variance
C'est ce que l'on appelle la variance totale.

b. matrice des correlations

Soit A : matrice des variables
A^s : matrice des variables centrées et reduites

On appelle matrice des corelations, la matrice $R = \frac{1}{n} A^s . A^s$

$$R = \begin{pmatrix} 1 & \rho(X,Y) & \rho(X,Z) \\ \rho(X,Y) & 1 & \rho(Y,Z) \\ \rho(X,Z) & \rho(Z,Y) & 1 \end{pmatrix}$$

Remarque : * R est carré et symetrique
* R passede une diagonale de 1
* La trace de R = nombre de variables