PARTIE I

1 ESPACE VECTORIEL

- **Addition de 2 vecteurs**
- **Multiplication par un scalaire**
- **Propriétés de l'addition**
- **Propriétés de la multiplication par un scalaire**
- **Propriétés déduites des précédentes**

1.1 Sous espace vectoriel

1.2 Combinaison linéaire

1.3 Indépendance linéaire

1.4 Base d'un espace vectoriel

PARTIE I

1.5 Longueur, distance, angle, produit scalaire

- Intérêt du produit scalaire
- Propriétés du produit scalaire

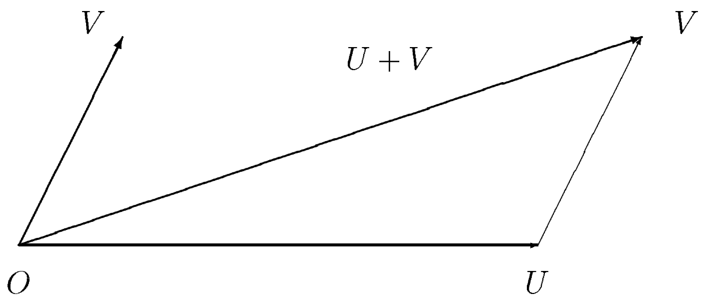
1.6 Décomposition en somme directe

2 PROJECTIONS

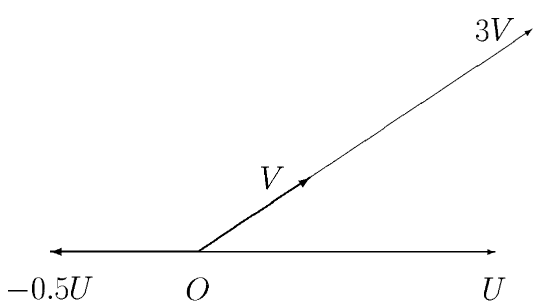
3 COORDONNEES DANS $\mathbb{R}^3$

1 ESPACE VECTORIEL

Addition de 2 vecteurs



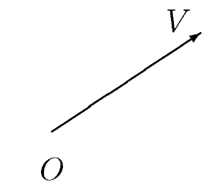
Multiplication par un scalaire



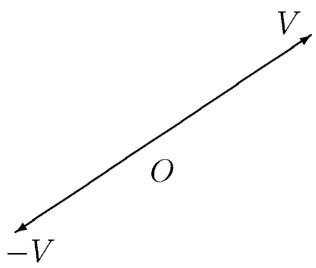
1 ESPACE VECTORIEL

1. On définit géométriquement un vecteur comme un segment de droite caractérisé par une longueur, une direction et un sens (*cf.* paragraphe 1.1).
2. On représente géométriquement l'addition de 2 vecteurs et la multiplication d'un vecteur par un scalaire.
3. On fait des digressions (au tableau) sur :
   - L'égalité de 2 vecteurs (*cf.* paragraphe 1.1.1).
   - La différence de 2 vecteurs (*cf.* paragraphe 1.1.3).
4. On donne la définition d'un espace vectoriel : collection d'éléments, appelés vecteurs, qui peuvent être additionnés et multipliés par des scalaires.

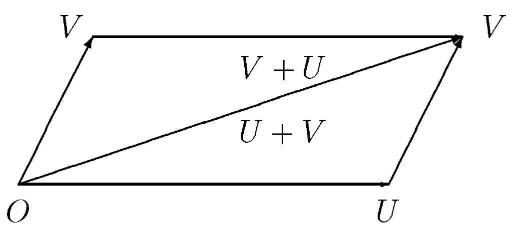
Propriétés de l'addition



$$\boxed{\phi + V = V + \phi = V}$$



$$\boxed{V + (-V) = \phi}$$



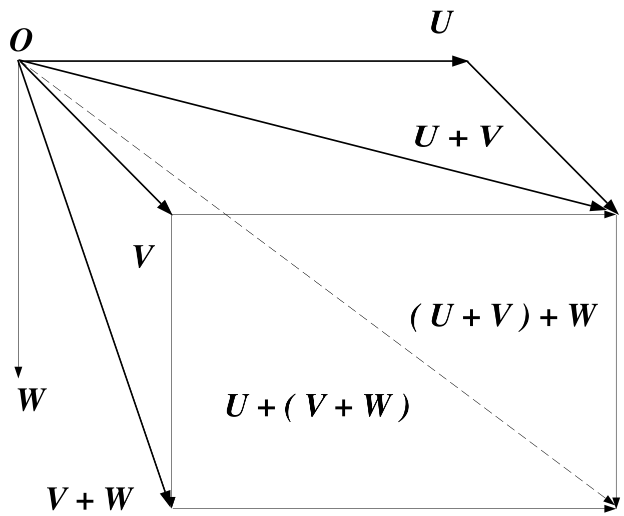
$$\boxed{U + V = V + U}$$

Propriétés de l'addition

Les propriétés de l'addition et de la multiplication par un scalaire sont données au paragraphe 3.1.

1. Existence d'un vecteur nul (ϕ).
2. Existence du vecteur $-V$ (l'opposé de V).
3. Commutativité de l'addition.

Propriétés de l'addition

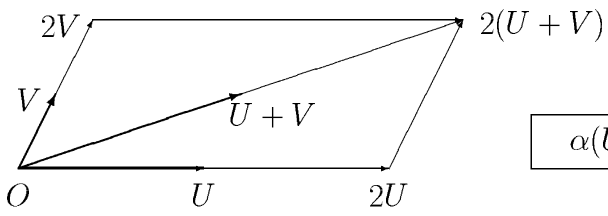


$$\boxed{U + (V + W) = (U + V) + W}$$

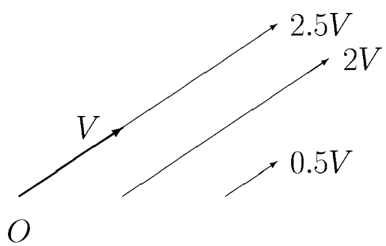
Propriétés de l'addition

1. Associativité de l'addition.

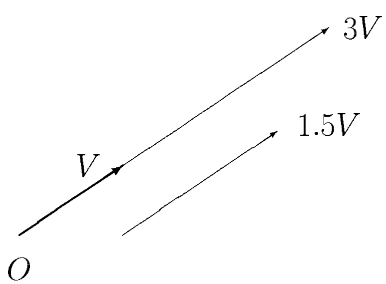
Propriétés de la multiplication par un scalaire



$$\boxed{\alpha(U+V) = \alpha U + \alpha V}$$



$$\boxed{(\alpha+\beta)V = \alpha V + \beta V}$$



$$\boxed{(\alpha\beta)V = \alpha(\beta V)}$$

Propriétés de la multiplication par un scalaire

1. Distributivité de la multiplication par rapport à l'addition :
   (a) Des vecteurs.
   (b) Des scalaires.
2. Associativité de la multiplication.

Propriétés déduites des précédentes

1. $(\alpha+\beta)U = \alpha U + \beta U$

$$si \;\; \beta = -\alpha \quad \Longrightarrow \begin{cases} (\alpha+\beta)U = o.U \\ \alpha U + \beta U = \alpha U + (-\alpha U) = \phi \end{cases}$$

$$\Downarrow$$

$$\boxed{o.U = \phi}$$

2. $\alpha(U+V) = \alpha U + \alpha V$

$$si \;\; V = -U \quad \Longrightarrow \begin{cases} \alpha(U+V) = \alpha\phi \\ \alpha U + \alpha V = \alpha U + (-\alpha U) = \phi \end{cases}$$

$$\Downarrow$$

$$\boxed{\alpha\phi = \phi}$$

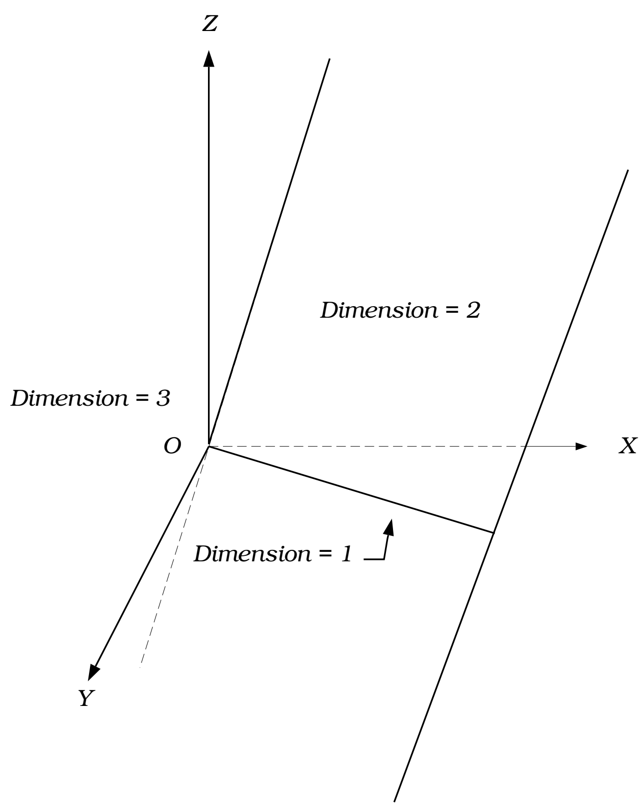
3. $\boxed{(-1)V = -V}$

Propriétés déduites des précédentes
(*cf.* paragraphe 3.1)

L'objectif est de montrer qu'il est possible de déduire des règles de travail à partir des propriétés de l'addition et de la multiplication par un scalaire. On montre que :

1. La multiplication d'un vecteur par zéro donne le vecteur nul.
2. La multiplication d'un scalaire par le vecteur nul donne le vecteur nul.
3. La multiplication d'un vecteur par -1 donne le vecteur opposé.

1.1 Sous espace vectoriel



1.1 Sous espace vectoriel

Un sous espace vectoriel est un sous ensemble de vecteurs d'un espace vectoriel, tel que les opérations (addition, multiplication par un scalaire) sur les vecteurs du sous espace donnent des vecteurs appartenant eux-mêmes au sous espace vectoriel.

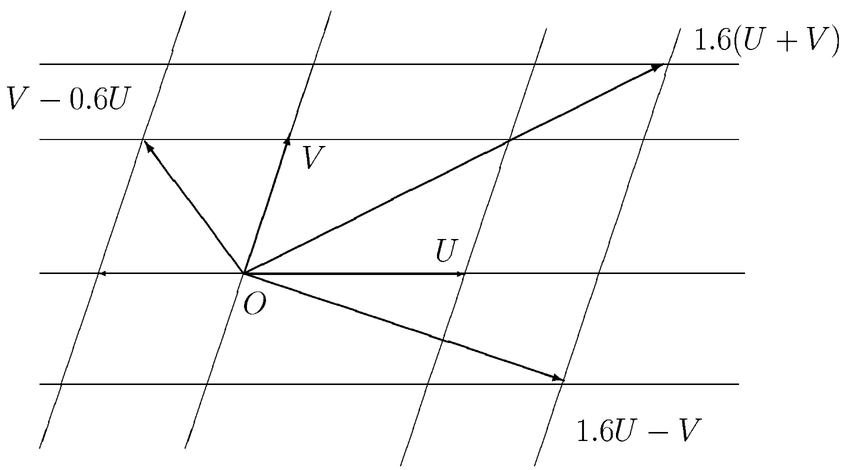
Exemples :

1. Le sous espace dont le seul élément est ϕ.
2. Une droite passant par l'origine est un sous espace vectoriel du plan.
3. Un plan contenant l'origine est un sous espace vectoriel de l'espace à trois dimensions.

Remarque :

Un sous espace qui ne contient pas l'origine n'est pas un sous espace vectoriel. C'est un sous espace **affine** (*cf.* paragraphe 3.7).

1.2 Combinaison linéaire



1.2 Combinaison linéaire

1. Soient U et V deux vecteurs, non colinéaires, du plan.

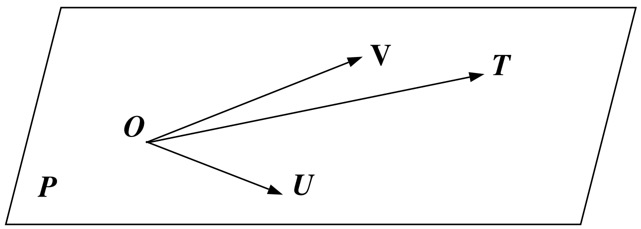
   On indique que les vecteurs $V - 0.6U$, $1.6(U+V)$ et $1.6U - V$, construits à partir des opérations de multiplication par un scalaire et d'addition, sont des combinaisons linéaires des vecteurs U et V.

2. On donne la définition d'une combinaison linéaire (*cf.* paragraphe 3.3) :

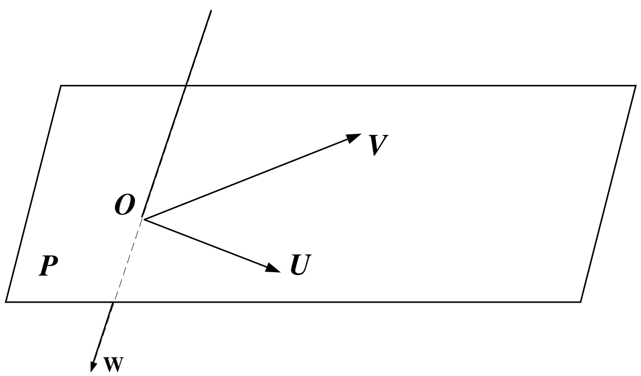
   V est une combinaison linéaire des vecteurs $U_1, U_2, \cdots, U_n$ s'il existe des scalaires $\alpha_1, \alpha_2, \cdots, \alpha_n$ tels que :

$$V = \alpha_1 U_1 + \alpha_2 U_2 + ... + \alpha_n U_n$$

1.3 Indépendance linéaire



Ces vecteurs sont-ils indépendants ?



U, V et W sont-ils indépendants ?

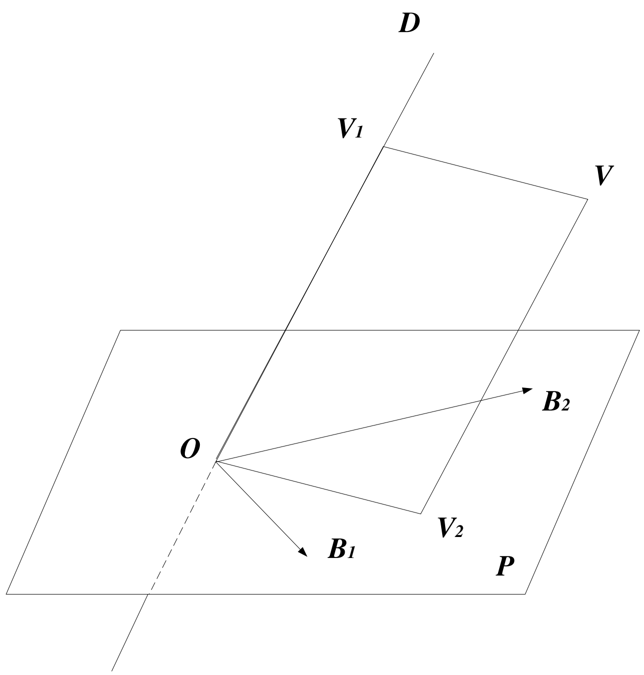
1.3 Indépendance linéaire

On définit l'indépendance linéaire (*cf.* paragraphe 3.4) :

Aucun vecteur ne peut être exprimé comme une combinaison linéaire des autres vecteurs.

1. Dans le premier cas de figure, chaque vecteur représenté dans le plan peut s'exprimer comme une combinaison linéaire des deux autres. Ces vecteurs sont linéairement dépendants.
2. Dans le deuxième cas de figure, aucun vecteur ne peut être représenté comme une combinaison linéaire des deux autres, ils sont linéairement indépendants.

1.6 Décomposition en somme directe



1.6 Décomposition en somme directe

(*cf.* paragraphe 3.8)

Soit l'espace vectoriel $\mathbb{R}^3$.

Cet espace peut être considéré comme la résultante des deux sous espaces vectoriels suivants : la droite D et le plan P.

Dans ce cas, on écrit que : $\mathbb{R}^3 = D \oplus P$

On dit que $\mathbb{R}^3$ est **somme directe** de D et de P.

On dit aussi que les sous espaces D et P sont **supplémentaires**.

On aurait aussi pu écrire que ; $\mathbb{R}^3 = D \oplus D_1 \oplus D_2$

avec D_1 et D_2 deux sous espaces vectoriels engendrés respectivement par B_1 et B_2 (B_1 et B_2 constituent une base) de P.

Ces notions se généralisent à plus de 3 dimensions.

1.4 Base d'un espace vectoriel de dimension finie

1. Une base d'un espace vectoriel est un ensemble de vecteurs linéairement indépendants qui engendrent l'espace dans sa totalité.
2. Tout vecteur est défini de façon unique comme combinaison linéaire des vecteurs d'une base.
3. Un espace vectoriel a toujours une base (mais elle n'est pas unique).
4. Chaque base a le même nombre d'éléments, ce nombre représente la dimension de l'espace.

1.4 Base d'un espace vectoriel de dimension finie

1. **Exemple :** deux vecteurs U et V non colinéaires forment une base du plan.
2. On montre (au tableau) que tout vecteur est défini de façon unique comme une combinaison linéaire des vecteurs de base.
   (a) Soit (U, V) une base du plan P.
   (b) On suppose que : $T = \alpha U + \beta V = \alpha' U + \beta' V$
   (c) On déduit que :

$$\begin{aligned} T - T &= \alpha U + \beta V - \alpha' U - \beta' V \\ &= (\alpha - \alpha')U + (\beta - \beta')V \\ &= \phi \end{aligned}$$

$$U \text{ et } V \text{ indépendants} \Longrightarrow \begin{cases} \alpha = \alpha' \\ \beta = \beta' \end{cases}$$

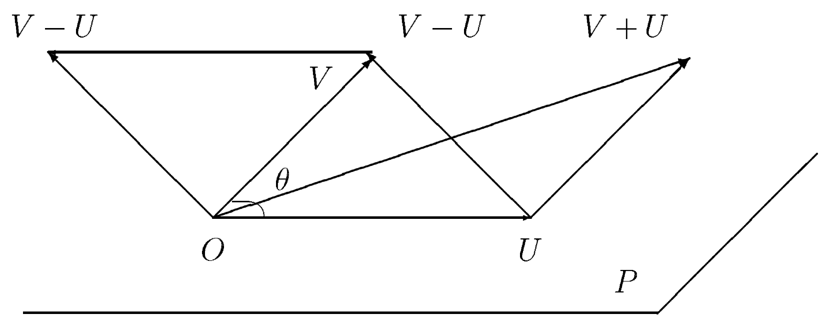
Définition du produit scalaire

1. $< \, , \, >$ application de $E \times E$ dans $I\!R$
2. $< U, V > = < V, U >$
3. $< U, U > \; \geq 0$
4. $< U, U > = 0$ si et seulement si $U = 0$
5. $< \alpha U + \beta V, W > = \alpha < U, W > + \beta < V, W >$

Définition du produit scalaire

La Définition du produit scalaire est énoncée au paragraphe 1.1.5.

1.5 Longueur, distance, angle, produit scalaire



1.5 Longueur, distance, angle, produit scalaire

(*cf.* paragraphe 1.1.5)

1. **Notations** :
   La longueur du vecteur U est notée $|U|$
   La distance de U à V est notée $|V-U|$
2. Soit θ l'angle entre U et V. Soit le triangle de sommets O, U et V, la règle du cosinus permet d'écrire :
   $$|U+V|^2 = |U|^2 + |V|^2 + 2|U||V|cos\theta \quad (1)$$
   Si l'angle θ est droit :
   $$|U+V|^2 = |U|^2 + |V|^2$$
3. La liaison entre U et V est décrite par $2|U||V|cos\theta$.
   $$\boxed{< U, V >= |U||V|cos\theta}$$
   $$(1) \Longleftrightarrow |U+V|^2 \; = |U|^2 \; |V|^2 + \; 2 < U, V >$$

Intérêt du produit scalaire

1. Longueur :
$$|U|^2 = < U, U >$$

2. Distance :
$$|V - U|^2 = < U, U > + < V, V > -2 < U, V >$$

3. Angle ($|U|, |V| \neq 0$) :
$$cos\theta = \frac{< U, V >}{|U||V|} = \frac{< U, V >}{(< U, U >< V, V >)^{1/2}}$$

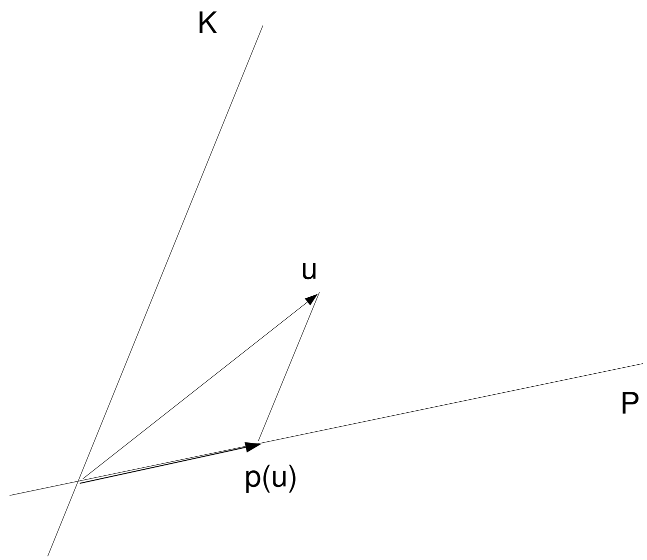
4. Orthogonalité :

   U et V sont orthogonaux si et seulement si :
$$< U, V > = 0$$

Intérêt du produit scalaire

On montre que le produit scalaire permet de calculer des longueurs, des distances, des angles. Ces propriétés sont énoncées au paragraphe 1.1.5.

2 PROJECTIONS



2 PROJECTIONS

(*cf.* paragraphe 1.3)

1. Définition d'une application linéaire

   f est une **application linéaire** de E dans E si pour tout élément U et V de E et tout élément α et β de $I\!R$, on a :

   $$f(\alpha U + \beta V) = \alpha f(U) + \beta f(V)$$

2. Définition d'une projection

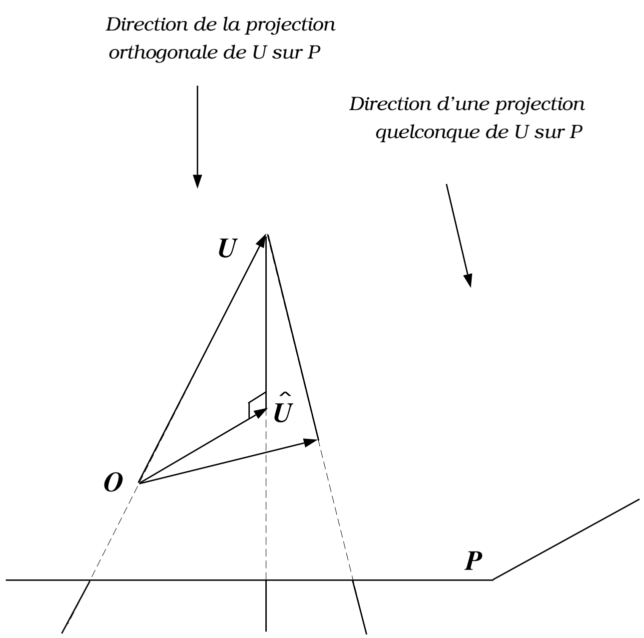
   Une **projection** p est une application linéaire vérifiant

   $$pop = p$$

3. Propriétés des projections

   Il existe un sous espace vectoriel P et K tel que :
   $E = P \oplus K$
   $\forall x \in K, \quad p(x) = 0$
   $\forall x \in P, \quad p(x) = x$
   On dit que p est la **projection sur P parallèlement à K**.

PROJECTION ORTHOGONALE



PROJECTION ORTHOGONALE

(*cf.* paragraphe 1.3)

1. Définition

   Deux **espaces vectoriels** E **et** E' **sont orthogonaux** si et seulement si : $\forall x \in E, \forall y \in E' \ \ < x, y >= 0$

2. Définition d'une projection orthogonale

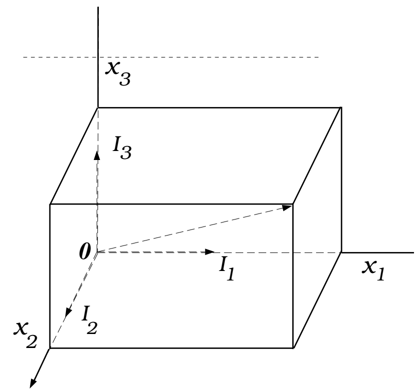
   On dit que p, **projection sur** P **parallèlement à** K **est orthogonale** si et seulement si $P \perp K$.

3. Propriétés des projections orthogonales

   Il existe un sous espace vectoriel P et K tel que :
   $< p(x), x - p(x) >= 0 \quad \forall x \in E$
   $|x|^2 = |p(x)|^2 + |x - p(x)|^2 \quad \forall x \in E$
   $\forall x \in E, \forall y \in P \quad \|x - p(x)\| \leq \|x - y|$

3 BASE ORTHONORMEE DE $I\!R^3$

Base orthonormée : $\begin{cases} |I_1| = |I_2| = |I_3| = 1 \\ < I_1, I_2 > = < I_1, I_3 > = < I_2, I_3 > = 0 \end{cases}$



3 BASE ORTHONORMEE DE $I\!R^3$

($cf.$ paragraphe 1.4)

On considère 3 vecteurs unitaires :

$$I_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad I_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad I_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

1. Ces vecteurs sont orthonormés :

$$< I_1, I_2 >=< I_1, I_3 >=< I_2, I_3 >= 0$$

$$< I_1, I_1 >=< I_2, I_2 >=< I_3, I_3 >= 1$$

2. Ils constituent une base. Tout vecteur X peut s'écrire :

$$X = x_1 I_1 + x_2 I_2 + x_3 I_3$$

3. x_1, x_2 et x_3 sont les coordonnées de X pour le système d'axes défini par I_1, I_2 et I_3.

Ces notions se généralisent à plus de 3 dimensions.

PARTIE II

GEOMETRIE ET STATISTIQUE

1 REPRESENTATION DES DONNEES

Les espaces des individus et des variables

2 LES RESUMES STATISTIQUES

Moyenne et variance

3 LES LIAISONS STATISTIQUES

Exemple : la liaison pression - température
Notion de corrélation

3 LES TESTS STATISTIQUES

Comparaison de deux populations

4 RESUME

1 REPRESENTATION DES DONNEES

Exemple :

Un tableau de données
les lignes $\Longleftrightarrow$ les individus
les colonnes $\Longleftrightarrow$ les variables

	var.1	*var*.2
ind.1	2	4
ind.2	4	3
ind.3	6	5

Que peut-on représenter géométriquement ?

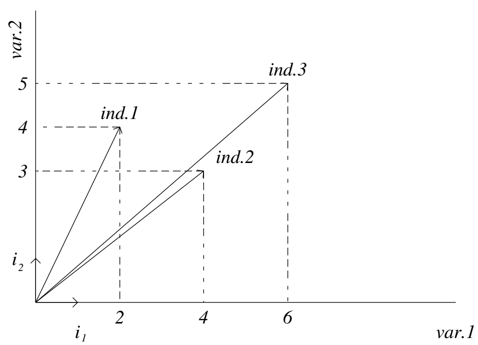
1 REPRESENTATION DES DONNEES
(*cf.* paragraphe 1)

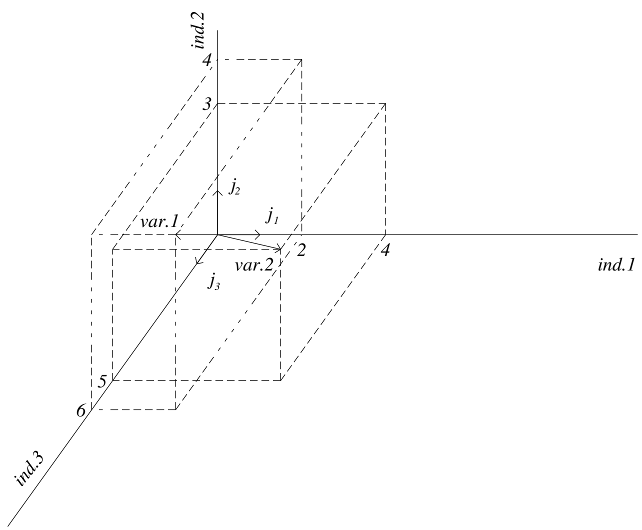
On considère le tableau de chiffres ci-contre. Il a 3 lignes (les individus) et 2 colonnes (les variables).

1. Les lignes et les colonnes du tableau sont des **vecteurs** lignes et colonnes qui peuvent être représentés géométriquement.
2. Pour appréhender toute l'information contenue dans le tableau, on lit le tableau ligne après ligne ou colonne après colonne. Quelque soit le mode de lecture adopté (ligne ou colonne), on lit toujours la totalité de l'information.

Il y a donc deux espaces vectoriels différents dans lesquels les données peuvent être représentées. Ce sont **les espaces** des **individus** et des **variables**.

L'espace des individus et l'espace des variables





L'espace des individus et l'espace des variables

1. Chaque ligne du tableau contient deux chiffres qui représentent les coordonnées de l'extrémité d'un **vecteur "individu"** dans l'espace défini par les deux variables. Cet espace est appelé **espace des individus**.
2. Chaque colonne du tableau contient trois chiffres qui représentent les coordonnées de l'extrémité d'un **vecteur "variable"** dans l'espace défini par les trois individus. Cet espace est appelé **espace des variables**.
3. Ces espaces ont des dimensions différentes (2 pour l'espace des individus et 3 pour l'espace des variables), ce sont des représentations différentes de l'information contenue dans le tableau des données. Ces deux espaces sont dits **espaces duaux**.

2 DES RESUMES STATISTIQUES

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$ une variable observée sur n individus.

$$J = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$ le vecteur à n composantes égales à 1.

$\widehat{X}_J$ la projection orthogonale de X sur J.

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ la moyenne.

$|X|^2$ le carré de la longueur du vecteur X

2 DES RESUMES STATISTIQUES

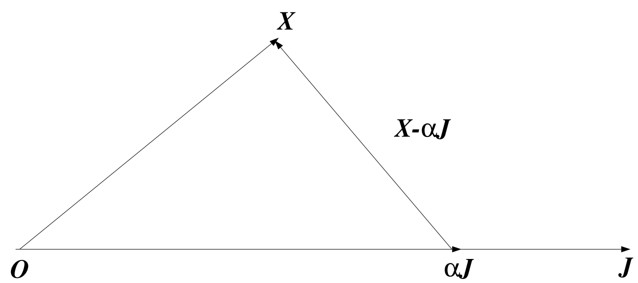
(*cf.* paragraphe 2)

Pour présenter les notions de moyenne, de variance, de coefficient de corrélation et de test, on se place dans l'espace des variables. Le but poursuivi est de :

montrer que l'on peut représenter des variables statistiques telles que des sommes de carrés ou des coefficients de corrélation.

On donne ici les principales notations qui seront utilisées par la suite.

Moyenne et variance



Moyenne et variance

Comment résumer l'information contenue dans le vecteur X ?

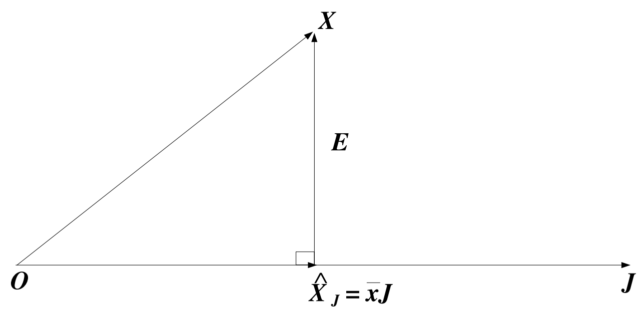
Une réponse à la question posée est :

En trouvant une image de X dans un sous espace vectoriel de dimension inférieure à n.

- Le résumé le plus simple est obtenu en projetant le vecteur X sur un espace à une dimension (une droite).

  On projette donc X sur J. Soit αJ un projeté quelconque de X sur J.
- Parmi tous les vecteurs αJ en existe-t-il un qui soit meilleur que tous les autres ?

Moyenne et variance



Moyenne et variance

Le meilleur résumé est celui pour lequel αJ est le plus proche de X.

- Soit $\widehat{X_J}$ le projeté orthogonal de X sur J. C'est le meilleur résumé car c'est pour $\widehat{X_J}$ que la distance $|X - \alpha J|$ est minimale. On a :

$$< X - \widehat{X_J}, J >= 0 \Longleftrightarrow < X - \alpha J, J >= 0$$

$$< X, J >= \alpha < J, J > \Longleftrightarrow \sum_{i=1}^{n} x_i = \alpha n$$

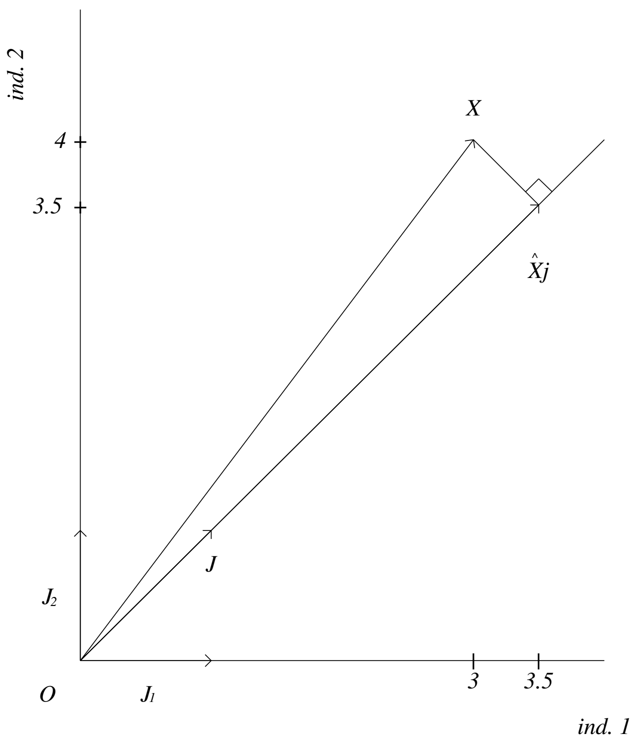
$$\alpha = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$$

  $\widehat{X_J}$ est donc le vecteur dont toutes les composantes α sont égales à la moyenne d'échantillonnage.

- On a par ailleurs $X = \bar{x}J + (X - \bar{x}J)$ avec $\bar{x}J$ dans un sous espace de dimension 1 et $E = X - \bar{x}J$ dans un sous espace de dimension $(n-1)$ appelé supplémentaire orthogonal à J.

Comment évaluer la qualité du résumé?

Moyenne et variance
Exemple : $X' = [3\ 4]$. Moyenne ?



Moyenne et variance

Cet exemple a pour objectif d'illustrer numériquement les résultats précédents.

$$|X|^2 =< X, X >= 4^2 + 3^2 = 25$$

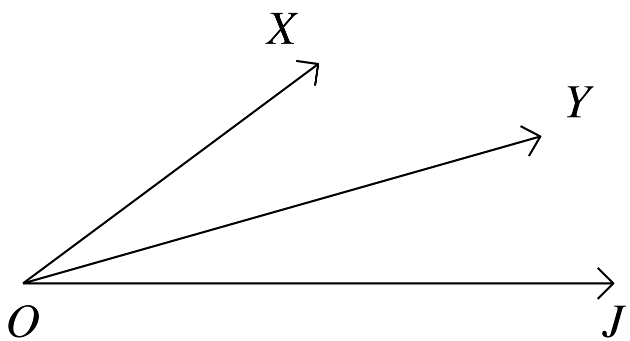
$$\widehat{X_J} = \begin{bmatrix} 3.5 \\ 3.5 \end{bmatrix}$$

$$|\widehat{X_J}|^2 =< \widehat{X_J}, \widehat{X_J} >= (3.5)^2 + (3.5)^2 = 24.5$$

$$|X - \widehat{X_J}|^2 =< X - \widehat{X_J}, X - \widehat{X_J} >$$

$$|X - \widehat{X_J}|^2 = (3 - 3.5)^2 + (4 - 3.5)^2 = 0.5$$

$$|\widehat{X_J}|^2 + |X - \widehat{X_J}|^2 = 24.5 + 0.5 = 25 = |X|^2$$



3 LES LIAISONS STATISTIQUES

1. On observe les variables X et Y sur n individus.
2. Comment peut-on mesurer l'intensité de la relation entre X et Y ?
3. Intuitivement, on peut penser à l'angle entre les vecteurs X et Y.

   Cet angle ne convient pas, on va le montrer sur l'exemple des pressions et des températures.

3 LES LIAISONS STATISTIQUES

- $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ $\quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ $\quad Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$ trois variables
- $\widehat{X}_J$, $\widehat{Y}_J$ et $\widehat{Z}_J$ les projections orthogonales de X, Y et Z sur J.
- X', Y' et Z' les variables X, Y et Z centrées.

Exemple : Soit le tableau suivant des pressions et des températures (degrés centigrade et Fahrenheit).

	P	T_c	T_f
ind.1	10	5	41
ind.2	20	15	59
ind.3	35	20	69

Quel angle exprime la liaison entre la pression et la température ?

3 LES LIAISONS STATISTIQUES

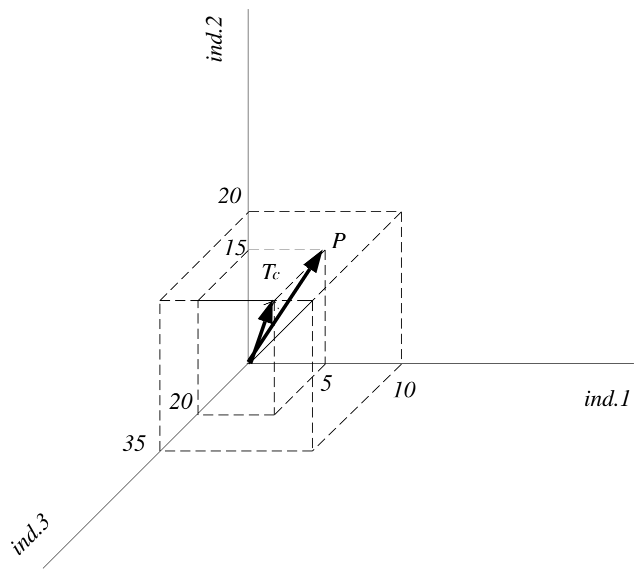
(*cf.* paragraphe 2.1.2)

1. On précise les notations qui vont être utilisées. On rappelle ce qu'est une variable centrée :

$$X' = X - \widehat{X}_J,\; Y' = Y - \widehat{Y}_J \text{ et } Z' = Z - \widehat{Z}_J$$

2. On présente les données de l'exemple des pressions et des températures.

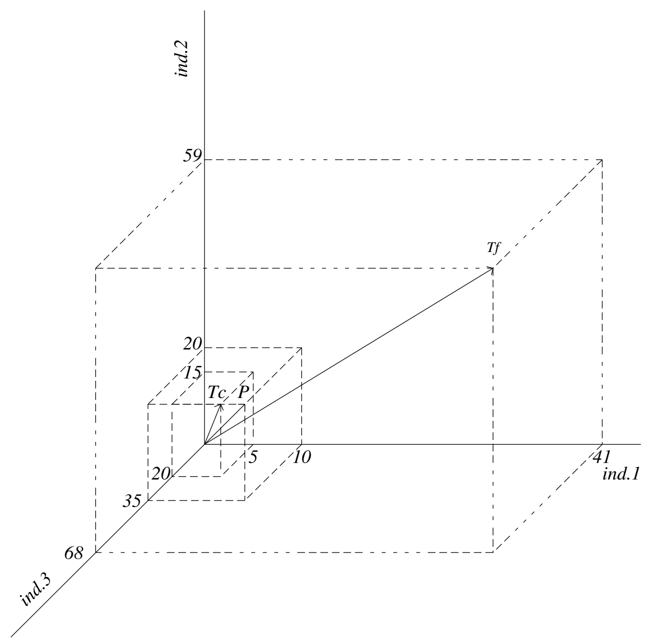
La liaison pression-température



La liaison pression-température

On représente le vecteur des pressions et le vecteur des températures exprimées en degrés centigrades.

La liaison pression-température



La liaison pression-température

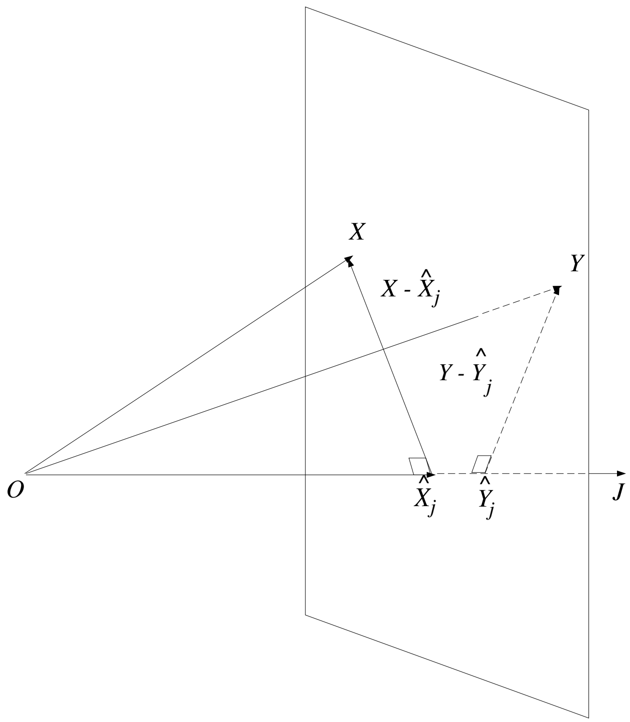
1. On représente le vecteur des pressions et le vecteur des températures exprimées en degrés Fahrenheit.
2. On fait remarquer que les deux vecteurs T_c et T_f ne sont pas colinéaires. Ils ne forment pas un même angle avec le vecteur P des pressions.

Donc, le changement d'unité de mesure de la température (une homothétie plus une translation) modifie la relation pression - température.

Les angles que font les vecteurs pression et températures ne sont pas de bonnes mesures de la liaison entre ces variables.

Dans ces conditions, que peut-on faire?

Notion de corrélation
L'espace des variables centrées



Notion de corrélation
L'espace des variables centrées

1. Existe-t-il un sous espace vectoriel dans lequel il est possible de trouver une mesure angulaire insensible à des translations et à des homothéties sur les variables?
2. La réponse est oui, c'est le sous espace des *variables centrées*.
3. On va le montrer sur l'exemple de la pression et des températures.

La liaison pression-température
L'espace des variables centrées

	P	T_c	T_f
$ind.1$	10	5	41
$ind.2$	20	15	59
$ind.3$	35	20	69

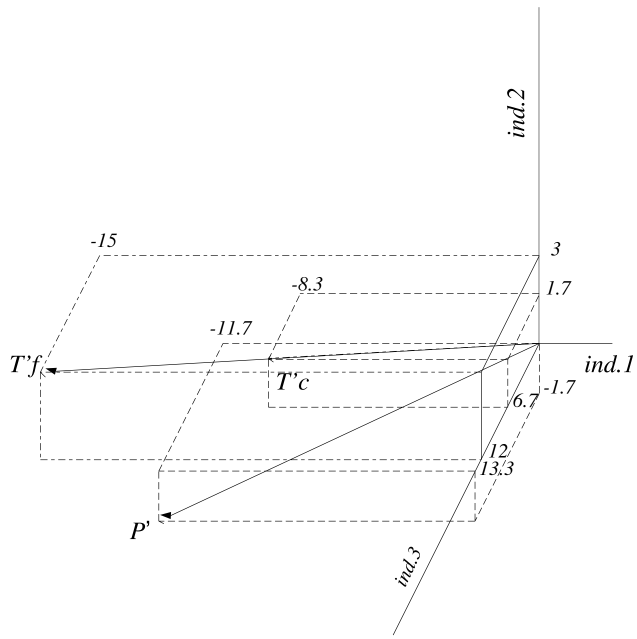
Centrons les variables

	P'	T'_c	T'_f
$ind.1$	-11.7	-8.4	-15
$ind.2$	-1.7	1.7	3
$ind.3$	13.4	6.7	12

La liaison pression-température
L'espace des variables centrées

1. On présente le tableau des données initiales et celui des variables centrées.
2. On va représenter sur un même graphique les variables centrées de la pression et des températures.

La liaison pression-température
L'espace des variables centrées

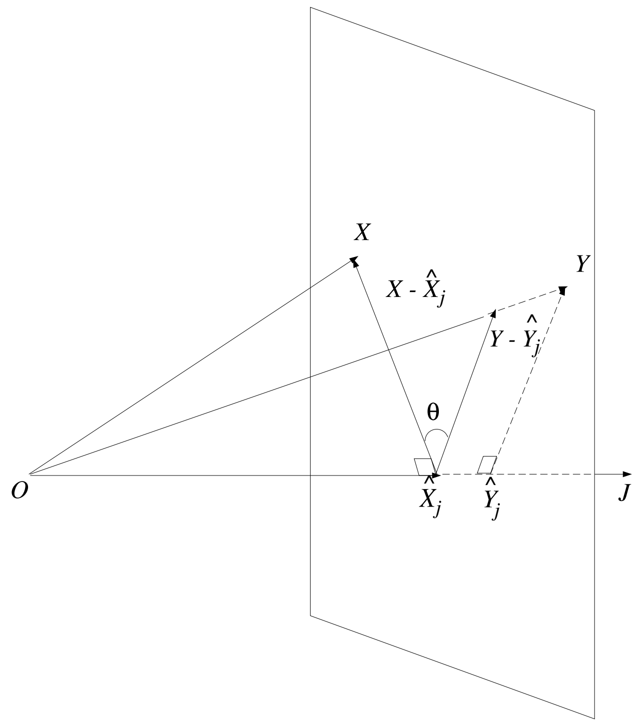


La liaison pression-température
L'espace des variables centrées

On observe que :

1. Les variables centrées des températures exprimées dans deux unités différentes sont représentées par des vecteurs colinéaires.
2. L'angle entre les variables centrées de la pression et des températures est insensible à une transformation sur les données initiales.

Notion de corrélation



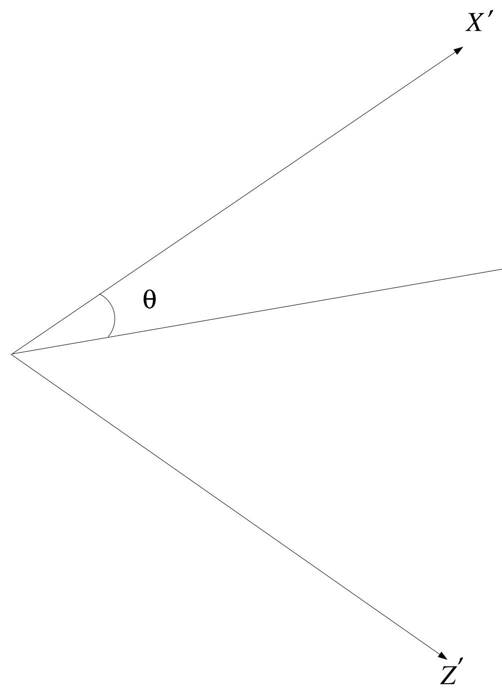
Notion de corrélation

1. On se place dans le sous espace orthogonal à J et on utilise comme mesure de l'intensité de la relation entre X et Y le cosinus de l'angle θ entre $X - \widehat{X_J}$ et $Y - \widehat{Y_J}$.
2. C'est cet angle θ que l'on a visualisé dans l'exemple pression-température, il est insensible à tout changement d'échelle et translation sur les X et les Y.
3. Le sous espace des variables centrées est le sous espace supplémentaire orthogonal à J.
4. On a :

$$cos(\theta) = \frac{< X - \widehat{X_J}, Y - \widehat{Y_J} >}{< X - \widehat{X_J}, X - \widehat{X_J} >< Y - \widehat{Y_J}, Y - \widehat{Y_J} >^{1/2}}$$

$$cos(\theta) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \rho(X,Y)$$

Notion de corrélation

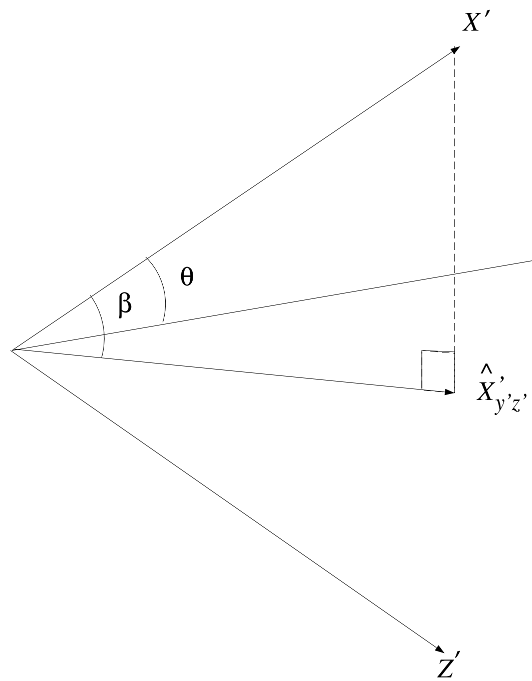


Notion de corrélation
(Corrélation simple)

On veut visualiser les différents coefficients de corrélation (simple, multiple et partielle). Pour cela :

- On choisit 3 variables X, Y et Z et on se place dans l'espace des variables centrées. Les vecteurs représentés sont X', Y' et Z'.
- Les angles que forment ces vecteurs entre eux ou avec des sous espaces vectoriels engendrés par ces vecteurs sont insensibles à des changements d'échelle et à des translations.
- Ici, on visualise θ l'angle dont le cosinus représente le coefficient de corrélation simple.

Notion de corrélation

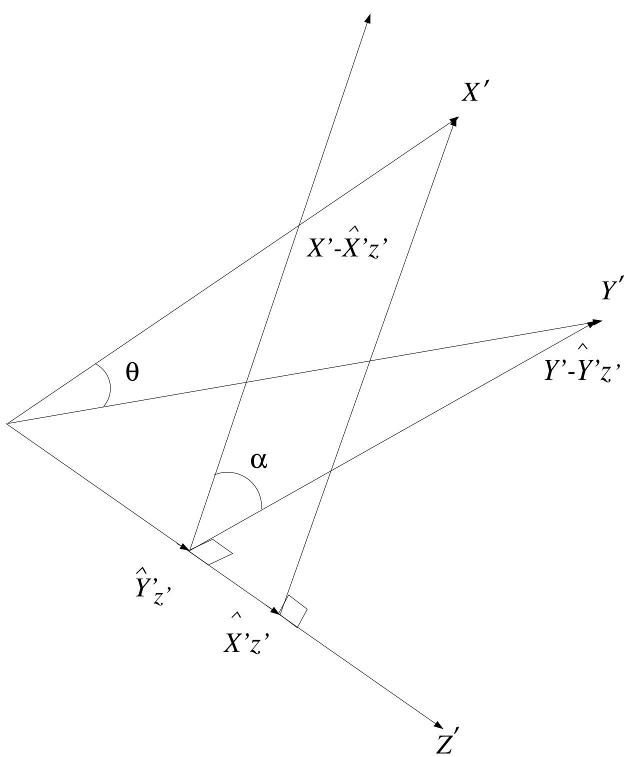


Notion de corrélation (Corrélation multiple)

Ici, on visualise le coefficient de corrélation simple et le coefficient de corrélation multiple.

- On note $\widehat{X'_{Y'Z'}}$ la projection orthogonale du vecteur X' sur le sous espace (un plan) engendré par les vecteurs Y' et Z'.
- Le cosinus de l'angle β entre X' et $\widehat{X'_{Y'Z'}}$ représente le coefficient de corrélation multiple entre X et le couple (Y, Z).

Notion de corrélation



INRA - Formation statistique - Module A-36

Notion de corrélation (corrélation partielle)

Ici, on visualise le coefficient de corrélation simple et le coefficient de corrélation partielle.

- On note $\widehat{X'_{Z'}}$ et $\widehat{Y'_{Z'}}$ les projections orthogonales de X' et Y' sur Z'.
- Par définition, le coefficient de corrélation partielle entre X et Y est le coefficient de corrélation entre X et Y lorsque l'effet de Z est éliminé.
- C'est donc le cosinus de l'angle α entre $X' - \widehat{X'_{Z'}}$ et $Y' - \widehat{Y'_{Z'}}$ qui représente le coefficient de corrélation partielle car ces vecteurs sont orthogonaux à Z' (leur corrélation avec Z' est nulle).

3 LES TESTS STATISTIQUES
Comparaison de deux populations

- P_1 et P_2 deux populations de moyenne μ_1 et μ_2
- Y le vecteur des observations (n_1 pour P_1 et n_2 pour P_2)
- J_1 et J_2 deux vecteurs orthogonaux tels que :

$$< J_1, J_2 >= 0 \text{ et } J = J_1 + J_2$$

$$Y = \begin{bmatrix} y_{11} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ \vdots \\ y_{2n_2} \end{bmatrix} \quad J_1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad J_2 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad J = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

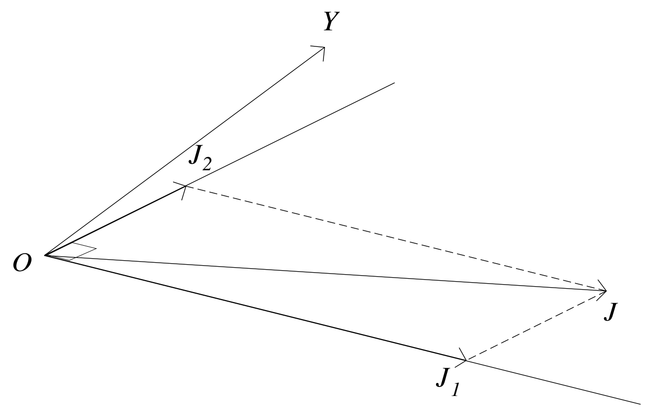
3 LES TESTS STATISTIQUES
Comparaison de deux populations

(*cf.* paragraphe 2.1.4)

Soient P_1 et P_2 deux populations normales de même variance σ^2 et de moyennes μ_1 et μ_2.

- On veut comparer ces deux populations c'est-à-dire répondre à la question : ces deux populations ont-elles des moyennes égales ($\mu_1 = \mu_2 = \mu$) ou ont-elles des moyennes différentes ($\mu_1 \neq \mu_2$) ?
- Pour répondre à cette question, on se place dans l'espace des variables de dimension $n = n_1 + n_2$.

Comparaison de deux populations

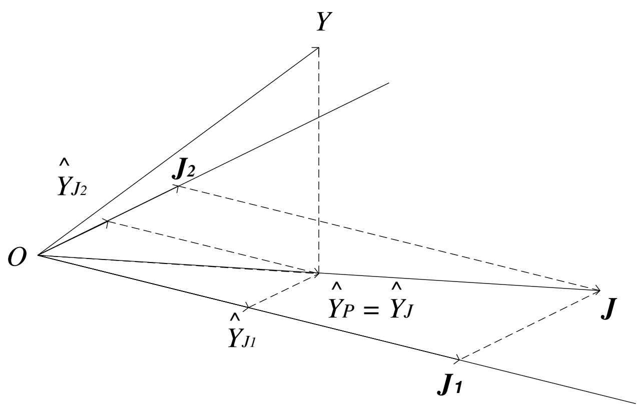


Comparaison de deux populations (Représentation des espaces)

- Pour chaque population (P_1 et P_2), les meilleurs résumés des observations sont les moyennes.
- Les moyennes sont les projections orthogonales de Y sur J_1 et J_2. Les sous espaces J_1 et J_2 sont orthogonaux ($< J_1, J_2 >= 0$).
- Le meilleur résumé de l'ensemble des données est la moyenne générale qui est la projection orthogonale de Y sur J.
- On se pose la question :

Les moyennes des deux populations sont-elles égales ?

Comparaison de deux populations (Les deux populations sont identiques)

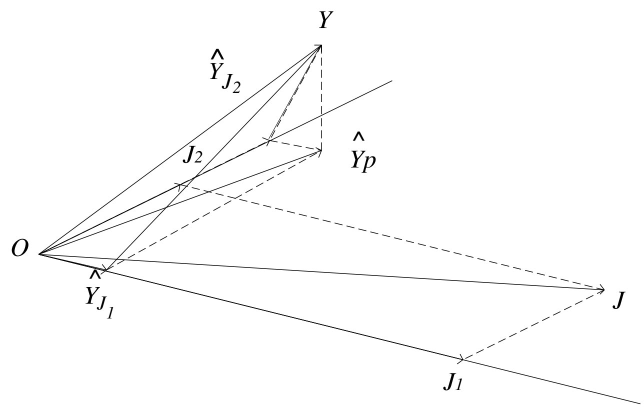


Comparaison de deux populations (Les deux populations sont identiques)

Si les deux populations P_1 et P_2 sont identiques :

- La moyenne des observations de P_1 est égale à la moyenne des observations de P_2 et ces deux moyennes sont égales à la moyenne calculée sur l'ensemble des observations.
- Sous cette condition, la projection orthogonale du vecteur Y des observations est $\widehat{Y_P} = \widehat{Y_J}$. Autrement dit, Y se projette sur J contenu dans P.

Comparaison de deux populations (Les deux populations sont différentes)



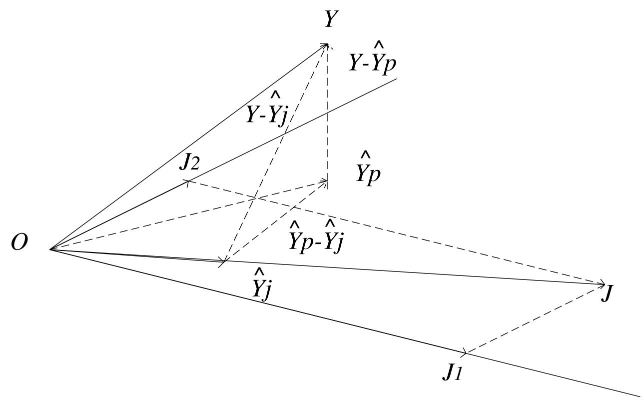
Comparaison de deux populations (Les deux populations sont différentes)

Si les deux populations P_1 et P_2 sont différentes :

- La moyenne des observations de P_1 est différente de la moyenne des observations de P_2 et ces deux moyennes sont différentes de la moyenne calculée sur l'ensemble des observations.
- Sous cette condition, la projection orthogonale du vecteur Y des observations est $\widehat{Y_P}$. Ici, Y ne se projette pas sur la bissectrice J de l'angle (J_1, J_2) contenu dans P.
- La distance de $\widehat{Y_P}$ à J sera d'autant plus grande que les deux moyennes seront différentes.
- On se pose la question :

Comment mesurer l'écart entre les moyennes des populations P_1 et P_2 ?

Comparaison de deux populations (Notion de test)



Comparaison de deux populations (Notion de test)

La différence entre les deux moyennes sera d'autant plus grande que $\widehat{Y_P}$ sera plus éloigné de J.

- L'éloignement est mesuré par la distance $|\widehat{Y_P} - \widehat{Y_J}|$. Le vecteur $\widehat{Y_P} - \widehat{Y_J}$ est dans un sous espace engendré par J_1 et J_2. Il est de dimension 1 et est orthogonal à J.
- La distance $|\widehat{Y_P} - \widehat{Y_J}|$ doit être appréciée par rapport à l'erreur $|Y - \widehat{Y_P}|$ que l'on commet quand on projette Y sur le sous espace engendré par J_1 et J_2. Le vecteur $Y - \widehat{Y_P}$ est dans l'orthogonal au sous espace engendré par J_1 et J_2, il a pour dimension $n_1 + n_2 - 2$.
- Pour apprécier la longueur du vecteur $\widehat{Y_P} - \widehat{Y_J}$ par rapport à celle de $Y - \widehat{Y_P}$ il faut tenir compte des dimensions des espaces qui les contiennent (les degrés de liberté).

$$\text{On forme le rapport } F = \frac{\frac{|\widehat{Y_P} - \widehat{Y_J}|^2}{1}}{\frac{|Y - \widehat{Y_J}|^2}{n_1 + n_2 - 2}}$$

4 Résumé

	Représentation géométrique	**Expression statistique**
Vecteur X	segment orienté de O vers X	$x_1, \cdots, x_n$
Vecteur J	segment orienté de O vers J	$1, \cdots, 1$
Vecteur $\widehat{X}_J = \bar{x}J$	Projeté orthogonal de X sur J	moyenne
$\lvert X - \widehat{X}_J \rvert^2$	Carré distance entre X et $\widehat{X}_J$	$(n-1)\times$ variance

4 Résumé

	Représentation géométrique	**Expression statistique**
$\dfrac{< X - \widehat{X}_J, Y - \widehat{Y}_J >}{\lvert X - \widehat{X}_J \rvert \lvert Y - \widehat{Y}_J \rvert}$	Cosinus angle $X - \widehat{X}_J$ et $Y - \widehat{Y}_J$	Corrélation simple entre X et Y
$\dfrac{\dfrac{\lvert \widehat{Y}_P - \widehat{Y}_J \rvert^2}{1}}{\dfrac{\lvert Y - \widehat{Y}_P \rvert^2}{n_1 + n_2 - 2}}$	Rapport carrés distances pondérées par d.d.l.	Fisher à 1 et $n_1 + n_2 - 2$ d.d.l. (si X et Y normales)

PARTIE III

1 MATRICES

- **Ranger des données**
- **Traiter des données**
- **Faire une transformation**

2 MATRICES PARTICULIERES

- **Matrice carrée**
- **Matrice nulle**
- **Matrice diagonale**
- **Matrice identité** I
- **Matrice** J
- **Matrice transposée :** A' **transposée de** A
- **Matrice symétrique**

PARTIE III

3 OPERATIONS MATRICIELLES

- **Addition**
- **Multiplication par un scalaire**
- **Multiplication de deux matrices**
- **Transposée d'un produit de matrices**
- **Trace**
- **Matrice idempotente**
- **Matrice inverse**
- **Notion de rang**
- **Déterminant d'une matrice**

1 MATRICES
Ranger des données

– **Exemple :** âge, poids et taille de 5 individus

$$\begin{bmatrix} 15 & 50 & 160 \\ 20 & 65 & 152 \\ 25 & 63 & 175 \\ 22 & 55 & 180 \\ 30 & 70 & 167 \end{bmatrix}$$

une ligne = un individu
une colonne = une variable

Matrice (5×3)

– **Notation**

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & a_{ij} & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} \qquad A_{(n\times p)} \; ou \; A_{(n,p)}$$

$$A = [a_{ij}]_{i=1,\cdots n \text{ et } j=1,\cdots p}$$

1 MATRICES
Ranger des données
(*cf.* paragraphe 1)

1. On montre sur un exemple que des observations faites sur des individus peuvent être rangées dans un tableau. Les lignes se réfèrent aux individus et les colonnes aux variables.
2. On rappelle que le tableau ainsi construit s'appelle une matrice et que les lignes et les colonnes du tableau sont des vecteurs lignes et des vecteurs colonnes.
3. On donne les notations couramment utilisées.
4. dimension = ordre d'une matrice

Traiter des données

– **Exemple :** âge, poids, taille de 5 femmes

$$F = \begin{bmatrix} 15 & 50 & 160 \\ 20 & 65 & 152 \\ 25 & 63 & 175 \\ 22 & 55 & 180 \\ 30 & 70 & 167 \end{bmatrix} \qquad F_{(5\times 3)}$$

– **Exemple :** âge, poids, taille de 5 hommes

$$H = \begin{bmatrix} 15 & 82 & 187 \\ 20 & 65 & 179 \\ 25 & 78 & 175 \\ 22 & 78 & 175 \\ 22 & 68 & 180 \\ 30 & 70 & 187 \end{bmatrix} \qquad H_{(5\times 3)}$$

– Rechercher les liaisons âge, poids, taille chez les femmes et chez les hommes.

Comparer les résultats.

Traiter des données

1. On prend l'exemple de trois variables observées sur deux groupes distincts d'individus, des femmes et des hommes. Les observations sont rangées dans deux matrices H et F. On précise les dimensions de ces matrices.
2. On dit que cette présentation des données est pratique pour calculer, indépendamment sur chaque groupe d'individus, des variables statistiques telles que des moyennes, des coefficients de corrélation, *etc.*
3. On aurait pu superposer les deux tableaux. Dans ce cas, les 5 premières lignes se seraient référées aux femmes et les 5 suivantes aux hommes. La matrice ainsi écrite aurait 10 lignes et 5 colonnes.

Faire une transformation

Exemple : Faire correspondre à chaque homme X d'âge (x_1), de poids (x_2) et de taille (x_3) :

- Un indice y_1 de risque de cancer.
- Un indice y_2 de risque d'infarctus du myocarde.

– Le produit de la matrice R des risques par unité d'âge, de poids et de taille par X permet d'obtenir le vecteur Y des risques y_1 et y_2.

$$Y_{(2\times1)} = R_{(2\times3)} \times X_{(3\times1)}$$

Faire une transformation

(*cf.* paragraphe 3.9.4)

1. E l'ensemble des hommes sur lesquels on a observé l'âge x_1, le poids x_2 et la taille x_3. Soit
$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$
2. F l'espace des couples Y avec y_1 le risque de cancer et y_2 le risque d'infarctus du myocarde.
$$Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$
3. r_{11}, r_{12}, r_{13} respectivement les risques de cancer associés à une unité d'âge, de poids et de taille.
r_{21}, r_{22}, r_{23} respectivement les risques d'infarctus du myocarde associés à une unité d'âge, de poids et de taille.
$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \end{bmatrix}$$

2 MATRICES PARTICULIERES

Matrice carrée : $M_{(n\times n)}$

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 3 & 0 & 1 \end{bmatrix}$$

Matrice nulle : $a_{ij} = 0 \; \forall \, i, \; j$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Matrice diagonale : $a_{ij} = 0$ si $i \neq j$

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2 MATRICES PARTICULIERES

(*cf.* paragraphe 4.1)

Matrice identité I **:** $a_{ii} = 1 \; \forall i$ et $a_{ij} = 0 \; \forall i \neq j$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Matrice J **:** $a_{ij} = 1 \; \forall \, i, \, j$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Matrice transposée : A' transposée de A $(a'_{ij} = a_{ji} \; , \; \forall \, i, \, j)$

$$A = \begin{bmatrix} 4 & 2 & 1 \\ -4 & -2 & -1 \end{bmatrix}$$

$$A' = \begin{bmatrix} 4 & -4 \\ 2 & -2 \\ 1 & -1 \end{bmatrix}$$

Matrice symétrique : $a_{ij} = a_{ji}$, $\forall\, i,\ j$

$$\begin{bmatrix} 4 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Une matrice carrée est symétrique si elle est égale à sa transposée ($A = A'$).

On n'écrit en général que la partie triangulaire haute ou basse. Soit :

$$\begin{bmatrix} 4 & 2 & 1 \\ & 2 & 1 \\ & & 1 \end{bmatrix}$$

$$\begin{bmatrix} 4 & & \\ 2 & 2 & \\ 1 & 1 & 1 \end{bmatrix}$$

3 OPERATIONS MATRICIELLES

Addition

Soient $A = [a_{ij}]$ et $B = [b_{ij}]$ des matrices de **même dimension**.
La somme de A et de B est une matrice C telle que :

$$C = A + B \text{ a pour éléments : } c_{ij} = (a_{ij} + b_{ij})$$

Exemple :

$$\begin{bmatrix} 1 & 0 \\ 5 & 4 \\ 8 & 4 \\ 3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 6 \\ 3 & 1 \\ 2 & 5 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1+1 & 0+6 \\ 5+3 & 4+1 \\ 8+2 & 4+5 \\ 3+2 & 2+2 \end{bmatrix}$$

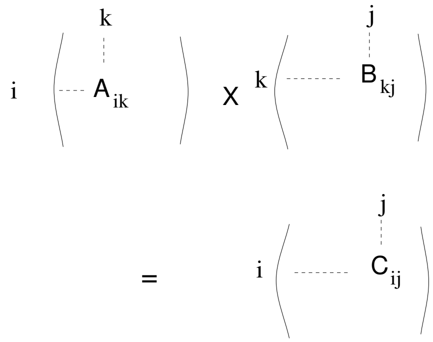
Multiplication par un scalaire

$$4 \times \begin{bmatrix} 1 & 0 \\ 5 & 4 \\ 8 & 2 \\ 3 & 2 \end{bmatrix} \quad = \quad ?$$

3 OPERATIONS MATRICIELLES

(*cf.* paragraphe 4.2)

Multiplication de deux matrices



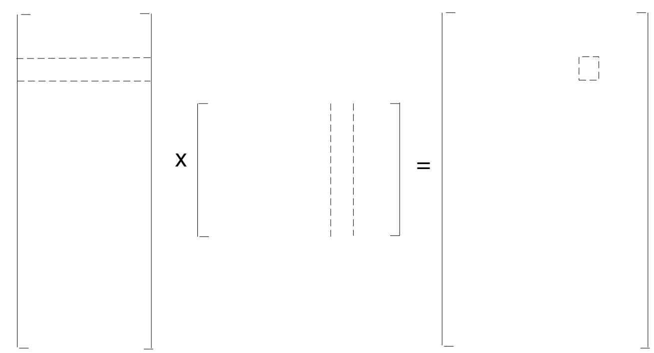
$A_{(n\times p)} \qquad B_{(p\times m)} \qquad \rightarrow \qquad C = A \times B$

Les éléments de $C_{(n\times m)}$ sont :

$$c_{ij} = \sum_{k=1}^{p} a_{ik}\, b_{kj}$$

Multiplication de deux matrices

1. Un autre dessin plus classique peut être fait au tableau, c'est le dessin suivant :



2. Multiplier ligne par colonne, c'est faire la somme des produits des termes de même rang dans la ligne et la colonne considérée.
3. Préciser la contrainte sur les dimensions : le nombre de colonnes de A doit être égal au nombre de lignes de B.

Multiplication de deux matrices

Exercices : effectuer les produits

$$\begin{array}{cc} & \begin{bmatrix}1&4&5&9&2\\0&0&1&2&3\\5&4&2&1&0\end{bmatrix} \\ \begin{bmatrix}1&3&0\\2&8&0\\0&2&1\end{bmatrix} & \begin{bmatrix}&&&&\\&&&&\\&&&&\end{bmatrix}\end{array}$$

$$\begin{array}{cc} & \begin{bmatrix}1&3&4&2&5\\8&2&0&0&0\\1&1&2&0&1\end{bmatrix} \\ \begin{bmatrix}2&0&1\\5&0&1\\9&0&2\\3&1&0\\0&2&1\end{bmatrix} & \begin{bmatrix}&&&&\\&&&&\\&&&&\\&&&&\\&&&&\end{bmatrix}\end{array}$$

Multiplication de deux matrices
Solutions

$$\begin{array}{cc} & \begin{bmatrix}1&4&5&9&2\\0&0&1&2&3\\5&4&2&1&0\end{bmatrix} \\ \begin{bmatrix}1&3&0\\2&8&0\\0&2&1\end{bmatrix} & \begin{bmatrix}1&4&8&15&11\\2&8&18&34&28\\5&4&4&5&6\end{bmatrix}\end{array}$$

$$\begin{array}{cc} & \begin{bmatrix}1&3&4&2&5\\8&2&0&0&0\\1&1&2&0&1\end{bmatrix} \\ \begin{bmatrix}2&0&1\\5&0&1\\9&0&2\\3&1&0\\0&2&1\end{bmatrix} & \begin{bmatrix}3&7&10&4&11\\6&12&22&10&26\\10&29&40&18&47\\11&11&12&6&15\\17&5&2&0&1\end{bmatrix}\end{array}$$

Le produit de deux matrices n'est pas en général commutatif.

Transposée d'un produit de matrices

$$(A.B)' = B'.A'$$

Exemple :

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & 1 \\ 2 & 1 \end{bmatrix}$$

$$(A.B) = \begin{bmatrix} 8 & 2 \\ 6 & 2 \end{bmatrix}$$

$$A' = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \qquad B' = \begin{bmatrix} 4 & 2 \\ 1 & 1 \end{bmatrix}$$

$$B'.A' = \begin{bmatrix} 8 & 6 \\ 2 & 2 \end{bmatrix} = (A.B)'$$

Trace

Donner la trace de $A = \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 8 & 3 & 0 \\ 4 & 9 & 2 & 4 \\ 2 & 6 & 1 & 5 \end{bmatrix}$

$$Tr(A) = \Sigma_{i=1}^{n} a_{ii}$$

Matrice idempotente

Une matrice $A_{(n \times n)}$ est idempotente si :

$$AA = A$$

Exemple : $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Trace

C'est la somme des éléments diagonaux de la matrice.

Matrice idempotente

L'idempotence est un concept important lié à la notion de projecteurs associés à une décomposition en somme directe (*cf.* paragraphe 4.5).

L'exemple correspond à la projection sur le premier vecteur de coordonnées (1,0).

Matrice inverse

On appelle matrice inverse de A la matrice A^{-1} telle que :

$$AA^{-1} = A^{-1}A = I$$

si A^{-1} existe, elle est unique, A est dite non singulière
si A^{-1} n'existe pas, A est dite singulière

Exemple : Inverse de $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$?

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \begin{cases} a &= 1 \\ b &= 0 \\ 2a + c &= 0 \\ 2b + d &= 1 \end{cases}$$

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Inverse d'une matrice diagonale?

Matrice inverse

(*cf.* paragraphe 4.2)

A est dite non singulière ou régulière ou inversible.

1. L'exemple traité ici pourra être repris après avoir vu le calcul d'un déterminant.

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$$

$$A^* = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \text{ (matrice adjointe)}$$

$$A^{*'} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \text{ (matrice adjointe transposée)}$$

$$det(A) = 1 \Longrightarrow A^{-1} = \frac{1}{det(A)}.A^{*'} = A^{*'}$$

2. L'inverse d'une matrice diagonale D est une matrice diagonale D^{-1} avec pour éléments diagonaux les inverses des éléments diagonaux de D.

NOTION DE RANG

Nombre de vecteurs linéairement indépendants

Vecteurs-ligne (si $n < p$)

Vecteurs-colonne (si $n > p$)

Si tous vecteurs sont linéairement indépendants alors la matrice est de plein rang.

Exemple :

$$A = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 20 \end{bmatrix} \qquad \begin{array}{c} Rang = 3 \\ \Longleftrightarrow \text{matrice} \\ \text{de plein rang} \end{array}$$

$$B = \begin{bmatrix} 10 & 10 & 10 \\ 15 & 15 & 15 \\ 3 & 6 & 9 \end{bmatrix} \qquad Rang = 2$$

NOTION DE RANG
($cf.$ paragraphe 4.2)

DETERMINANT D'UNE MATRICE

Soit le système d'équations :

$$\begin{cases} 3x_1 + 5x_2 + 1x_3 = 7 \\ 2x_1 + 4x_2 + 5x_3 = 4 \\ 1x_1 + 2x_2 + 2x_3 = 6 \end{cases}$$

on l'écrit matriciellement :

$$\begin{bmatrix} 3 & 5 & 1 \\ 2 & 4 & 5 \\ 1 & 2 & 2 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ 6 \end{bmatrix}$$

$$\Updownarrow$$

$$AX = B$$

Si A est inversible alors le système d'équations admet une solution :

$$X = A^{-1}B$$

Pour savoir si A^{-1} existe, on calcule son déterminant.

$$det(A) \neq 0 \Longleftrightarrow A^{-1} \text{ existe.}$$

DETERMINANT D'UNE MATRICE
(*cf.* paragraphe 4.2)

1. On pourra commencer par une matrice 2x2.
2. On donne au tableau la notation suivante :
$$det(A) = \begin{vmatrix} 3 & 5 & 1 \\ 2 & 4 & 5 \\ 1 & 2 & 2 \end{vmatrix}$$
3. On rappelle les définitions suivantes :
   - Le **mineur** de l'élément a_{ij} de A est le déterminant obtenu en supprimant la ligne i et la colonne j de A.
   - Le **cofacteur** de a_{ij} est le mineur multiplié par $(-1)^{i+j}$.
   - La matrice A^* des cofacteurs de A est appelée **matrice adjointe**.
4. On écrit (au tableau) le mineur de l'élément $a_{11} = 3$, le cofacteur de l'élément $a_{12} = 5$ puis, on en fait calculer un ou deux autres par les stagiaires.

Calcul du déterminant

Soit la matrice :

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\begin{aligned} det(A) = \; & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \end{aligned}$$

Exemple :

$$A = \begin{bmatrix} 3 & 5 & 1 \\ 2 & 4 & 5 \\ 1 & 2 & 2 \end{bmatrix}$$

$$det(A) = 24 + 25 + 4 - 4 - 30 - 20 = -1$$

$$A \text{ non singulière} \iff \begin{array}{l} \text{1- } A \text{ de plein rang} \\ \text{2- } det(A) \neq 0 \\ \text{3- } A^{-1} \text{ existe} \end{array}$$

Calcul du déterminant

1. On peut montrer (au tableau) le calcul du déterminant par la méthode des mineurs. Soit :

$$det(A) = 3\begin{vmatrix} 4 & 5 \\ 2 & 2 \end{vmatrix} - 5\begin{vmatrix} 2 & 5 \\ 1 & 2 \end{vmatrix} + 1\begin{vmatrix} 2 & 4 \\ 1 & 2 \end{vmatrix}$$

$$det(A) = 3 \times (-2) - 5 \times (-1) + 1 \times (0) = -1$$

2. On demande aux stagiaires de calculer le déterminant par la méthode des mineurs associés aux éléments d'une colonne.

PARTIE IV

VALEURS ET VECTEURS PROPRES

1 Notion de valeur et de vecteur propre

- **Exemple : Symétrie par rapport à la** *1ère* **bissectrice**
- **Généralisation**

2 Propriétés

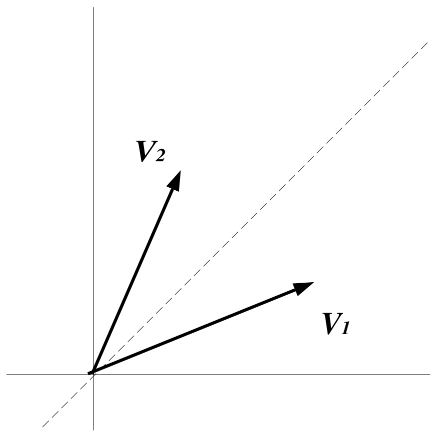
3 Exercices

- **Symétrie par rapport à la** *1ère* **bissectrice**

1 Notion de valeur et de vecteur propre
Symétrie par rapport à la *1ère* bissectrice

Soit l'application $f_1 \;:\; I\!R^2 \rightarrow I\!R^2$

$$f_1 \;:\; V_1 = \left[\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right] \longrightarrow V_2 = \left[\begin{smallmatrix} x_2 \\ y_2 \end{smallmatrix}\right]$$



La matrice A_1 associée à f_1 est $A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Quels sont les vecteurs V_1 non nuls du plan tels que :

$$A_1.V_1 \;=\; \lambda V_1 \text{ avec } \lambda \in I\!R \text{ ?}$$

1 Notion de valeur et de vecteur propre
Symétrie par rapport à la *1ère* bissectrice
(*cf.* paragraphe 4.4)

Soit f_1 l'application de $I\!R^2$ dans $I\!R^2$ qui à tout vecteur V_1 associe son symétrique par rapport à la 1*ère* bissectrice. Le transformé d'un vecteur V_1 de coordonnées $\left[\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right]$ est un vecteur V_2 de coordonnées $\left[\begin{smallmatrix} x_2 \\ y_2 \end{smallmatrix}\right]$ telles que $\left\{\begin{smallmatrix} x_2 = y_1 \\ y_2 = x_1 \end{smallmatrix}\right.$

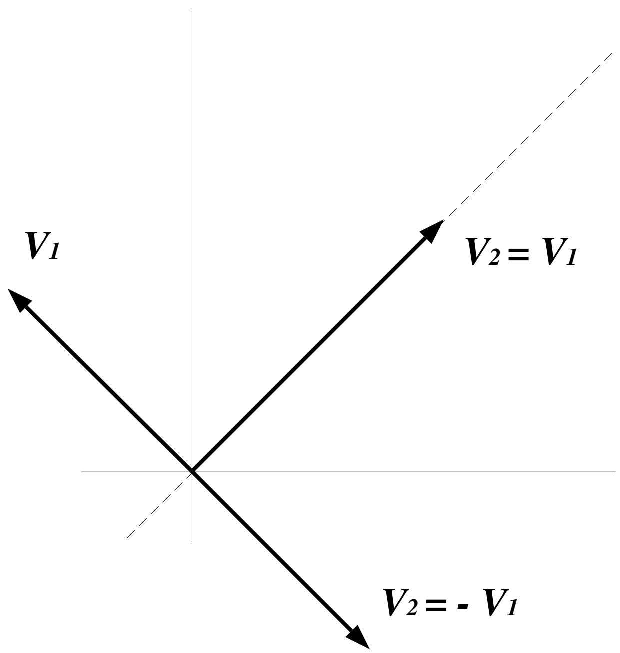
A cette application f_1 est associée une matrice :

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

1. On vérifiera au tableau qu'on a bien $V_2 = A_1.V_1$, c'est-à-dire $\left[\begin{smallmatrix} x_2 \\ y_2 \end{smallmatrix}\right] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \left[\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} y_1 \\ x_1 \end{smallmatrix}\right]$
2. On recherche l'ensemble des vecteurs V_1 non nuls du plan dont le transformé par f_1 est un vecteur colinéaire à V_2. Autrement dit tel que :

$$V_2 = \lambda V_1 \text{ avec } \lambda \in I\!R.$$

Symétrie par rapport à la *1ère* bissectrice



Symétrie par rapport à la *1ère* bissectrice

Deux ensembles de vecteurs satisfont à cette propriété :

1. L'ensemble des vecteurs $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ tels que eux-mêmes.
   Pour ces vecteurs, on a $\lambda = 1$.
2. L'ensemble des vecteurs $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ tels que $y_1 = -x_1$. Ces vecteurs sont transformés en un vecteur de sens opposé.
   Pour ces vecteurs, on a $\lambda = -1$.

Ces deux ensembles de vecteurs sont appelés **vecteurs propres** de l'application f_1 (ou de la matrice A_1) associés respectivement aux **valeurs propres** $\lambda = 1$ et $\lambda = -1$.

Généralisation

Soit :

- E un espace vectoriel de dimension n.
- A matrice.

$V \in E$, non nul, est un **vecteur propre** de A associé à la **valeur propre** $\lambda \in \mathbb{R}$ si et seulement si $A.V = \lambda V$.

Généralisation

2 Propriétés

1. A une matrice carrée A d'ordre n sont associées au plus n valeurs propres réelles.
2. Les valeurs propres sont les racines d'un polynôme de degré n appellé **polynôme caractéristique**. Elles satisfont à l'équation : $\det(A - \lambda I) = 0$ avec I = matrice identité.
3. Si A est une matrice **symétrique** (cas des matrices d'inertie et de variance-covariance) :
   - toutes les racines du polynôme caractéristique sont **réelles**.
   - Si toutes les racines sont positives la matrice A est dite **définie positive**.
   - Si λ_i et λ_j sont deux valeurs propres distinctes les vecteurs propres associés sont **orthogonaux**.

2 Propriétés

($cf.$ paragraphe 4.4.2)

Soit une matrice $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, λ et $V = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ une valeur et un vecteur propre de A.

Par définition $A.V = \lambda V$. Soit : $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda y_1 \end{bmatrix}$

d'où $\begin{cases} a_{11}x_1 + a_{12}y_1 = \lambda x_1 \\ a_{21}x_1 + a_{22}y_1 = \lambda y_1 \end{cases} \Leftrightarrow \begin{cases} (a_{11} - \lambda)x_1 + a_{12}y_1 = 0 \\ a_{21}x_1 + (a_{22} - \lambda)y_1 = 0 \end{cases}$

On a là un système homogène qui admet des solutions non nulles si et seulement si :

$$\det \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} = 0$$

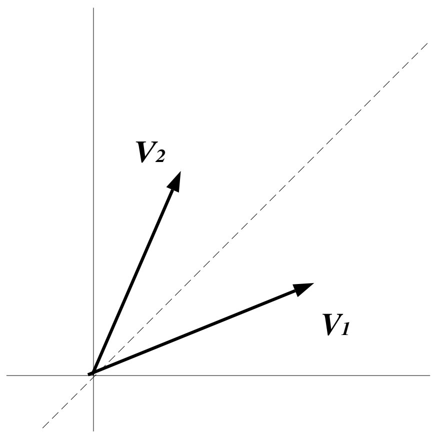
Les valeurs propres de A sont donc les solutions de l'équation

$$det(A - \lambda I) = 0.$$

3 Exercices
Symétrie par rapport à la *1ère* bissectrice

Soit l'application $f_1 \ : \ \mathbb{R}^2 \to \mathbb{R}^2$

$$V_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \longrightarrow V_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \text{avec} \begin{cases} x_2 = y_1 \\ y_2 = x_1 \end{cases}$$



La matrice A_1 associée à f_1 est $A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Rechercher les valeurs et les vecteurs propres de la matrice A_1 ?

Symétrie par rapport à la *1ère* bissectrice

(On donne la solution au tableau)

1. Les valeurs propres de la matrice A_1 sont les solutions de l'équation $det(A_1 - \lambda I) = 0$. Soit :

$$\det \begin{bmatrix} 0-\lambda & 1 \\ 1 & 0-\lambda \end{bmatrix} = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = 0$$

   Il y a donc 2 solutions $\lambda_1 = 1$ et $\lambda_2 = -1$.

2. Les vecteurs propres associés à la valeur propre $\lambda_1 = 1$ sont tels que :

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 \\ \lambda_1 y_1 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

   D'où $y_1 = x_1$;
   $e_1 \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ est un vecteur propre, c'est celui de norme 1.
   De même les vecteurs propres associés à la valeur propre $\lambda_2 = -1$ sont tels que $y_2 = -x_2$;
   $e_2 \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.
   On retrouve bien les valeurs et les vecteurs propres trouvés intuitivement.

   Les 2 vecteurs propres sont orthogonaux dans la base (e1,e2) et la matrice s'écrit : $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

PARTIE V

MAXIMISATION-DERIVATION

1 Recherche des extrema

- **Extremum d'une fonction**

2 Dérivation

- **Dérivée première**
- **Exemple**
- **Dérivée seconde**

3 Recherche des extrema

- **Exemples**

Extremum d'une fonction d'une variable

$$f : [a, b] \longrightarrow I\!R$$

1. dérivable deux fois.
2. Calcul de $f'(x)$
   f admet un extremum en $x_0 \in]a, b[\Longrightarrow \; f'(x_0) = 0.$
3. Calcul de $f"(x)$

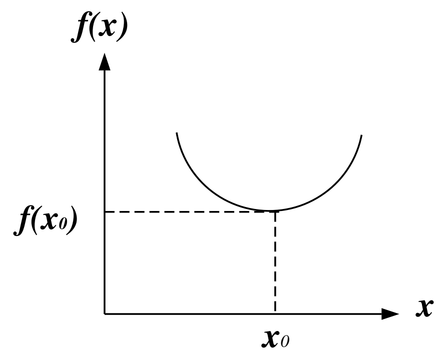
$$\begin{cases} f(x_0) \text{ est minimum si } f''(x_0) > 0 \\ f(x_0) \text{ est maximum si } f''(x_0) < 0 \end{cases}$$

Contre exemples :

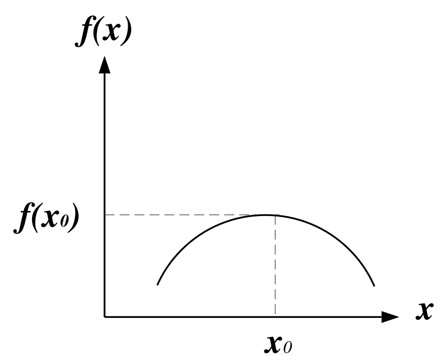
- $f(x) = x^3 \quad f'(0) = 0$, pas d'extremum en 0
- $f(x) = x^4 \quad f'(0) = 0 \quad f''(0) = 0$, minimum en 0
- $f(x) = x$ sur [0,1] , 0 est un minimum mais $f'(0) \neq 0$
- $f(x) = |x|$, 0 est un minimum mais f non dérivable en 0

Extremum d'une fonction d'une variable

f(x₀) est un minimum si *f ''(x₀) > 0*



f(x₀) est un maximum si *f ''(x₀)*



2 Dérivation
Dérivée première

$$f \;:\; \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \longrightarrow f(X)$$

$$\frac{\partial f(X)}{\partial X} = \begin{bmatrix} \frac{\partial f(X)}{\partial x_1} \\ \frac{\partial f(X)}{\partial x_2} \\ \vdots \\ \frac{\partial f(X)}{\partial x_n} \end{bmatrix}$$

2 Dérivation
Dérivée première
(*cf.* paragraphe 5.2)

1. Définition de la dérivée partielle $\frac{\partial f(X)}{\partial x_i}$.

   On fixe les x_j pour $j \neq i$ et on dérive en x_i comme d'habitude.
2. Définition du gradient : C'est le vecteur

$$\frac{\partial f(X)}{\partial X} = \begin{bmatrix} \frac{\partial f(X)}{\partial x_1} \\ \frac{\partial f(X)}{\partial x_2} \\ \vdots \\ \frac{\partial f(X)}{\partial x_n} \end{bmatrix}$$

Exemple

$$f \;:\; \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \longrightarrow f(X) = a_1 x_1^2 + a_2 \sin(x_2) + a_3$$

$$\frac{\partial f(X)}{\partial X} = \begin{bmatrix} \frac{\partial f(X)}{\partial x_1} \\ \frac{\partial f(X)}{\partial x_2} \end{bmatrix}$$

$$\frac{\partial f(X)}{\partial X} = \begin{bmatrix} 2a_1 x_1 \\ a_2 \cos(x_2) \end{bmatrix}$$

Quelques dérivées premières

Soient

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ et } B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

- Si pour tout X, $f(X) = B'.X = X'.B$ alors :
$$\frac{\partial f(X)}{\partial X} = B$$
- Si pour tout X, $f(X) = \text{Constante}$, alors :
$$\frac{\partial f(X)}{\partial X} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

On pourra effectuer le calcul explicitement.

Quelques dérivées premières

$$f(X) = \sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + \cdots + x_n^2$$

soit

$$f(X) = [x_1 x_2 \cdots x_n] \, . \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = X'.X$$

alors

$$\frac{\partial f(X)}{\partial X} = \frac{\partial (X'.X)}{\partial X} = 2.X$$

Quelques dérivées premières

$$f(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}.x_i.x_j$$

$$\begin{aligned} f(X) = \; & a_{11}.x_1^2 + a_{12}.x_1.x_2 + \cdots + a_{1n}.x_1.x_n \\ & + a_{21}.x_2.x_1 + a_{22}.x_2^2 + \cdots + a_{2n}.x_2.x_n \\ & + \cdots + a_{nn}.x_n^2 \end{aligned}$$

$$f(X) = [x_1 x_2 \cdots x_n] \, . \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}}_{A} . \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$f(X) = X'.A.X$$

Quelques dérivées premières

$$f(X) = X'.A.X$$

alors

$$\frac{\partial f(X)}{\partial X} = \frac{\partial (X'.A.X)}{\partial X} = A.X + A'.X$$

si de plus A est symétrique ($A = A'$) alors

$$\frac{\partial f(X)}{\partial X} = 2.A.X$$

Dérivée seconde

Quand $f(X)$ est une forme quadratique

$$f(X) = X'.A.X$$

$$\frac{\partial f(X)}{\partial X} = A.X + A'.X$$

alors

$$\frac{\partial^2 f(X)}{\partial X^2} = \frac{\partial(A.X + A'.X)}{\partial X} = A' + A$$

si de plus A est symétrique ($A = A'$)

$$\frac{\partial^2 f(X)}{\partial X^2} = 2.A$$

C'est une matrice appelée :
matrice **HESSIENNE** (H)

3 Recherche des extrema

$f \ : \ \mathbb{R}^n \longrightarrow \mathbb{R}$ 2 fois dérivable

vecteur gradient : $\begin{bmatrix} \vdots \\ \dfrac{\partial f(X)}{\partial x_i} \\ \vdots \end{bmatrix}$ et matrice hessienne : $\left[\dfrac{\partial^2 f(X)}{\partial x_i \partial x_j} \right]$

Condition nécessaire pour obtenir un extremum de $f(X)$ au point $X = X_0$.

$$\begin{bmatrix} \vdots \\ \dfrac{\partial f(X)}{\partial x_i} \\ \vdots \end{bmatrix}_{X=X_0} = 0$$

Condition suffisante pour avoir un :

1. **Maximum :** matrice hessienne évaluée en X_0 définie négative.
2. **Minimum :** matrice hessienne évaluée en X_0 définie positive.

Recherche des extrema
($cf.$ paragraphe 5.3)

1. On résout l'équation : $\dfrac{\partial f(X)}{\partial X} = 0$
2. La matrice hessienne est symétrique, elle peut être ni définie positive ni définie négative (point-selle).

Exemples

Déterminer les extrema s'ils existent :

1. Soit f une application de $\mathbb{R}^2$ dans $\mathbb{R}$:

$$f(x,y) = x^2 + y^2$$

2. Soit f une application de $\mathbb{R}^2$ dans $\mathbb{R}$:

$$f(x,y) = -x^2 + 2x - y^2 + 4y - 7$$

3. Soit f une application de $\mathbb{R}^2$ dans $\mathbb{R}$:

$$f(x,y) = x^2 - y^2$$

Exemple 1

Calcul du gradient :

$$\begin{bmatrix} \dfrac{\partial f(x,y)}{\partial x} \\ \dfrac{\partial f(x,y)}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

Recherche d'un extremum :

$$\begin{bmatrix} \dfrac{\partial f(x,y)}{\partial x} \\ \dfrac{\partial f(x,y)}{\partial y} \end{bmatrix} = 0 \Longleftrightarrow (x,y) = (0,0)$$

Calcul de la matrice hessienne :

$$\begin{bmatrix} \dfrac{\partial^2 f(x,y)}{\partial x^2} & \dfrac{\partial^2 f(x,y)}{\partial x \partial y} \\ \dfrac{\partial^2 f(x,y)}{\partial y \partial x} & \dfrac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

La matrice hessienne est définie positive, le point $(0,0)$ est un minimum.

Exemple 2

Calcul du gradient :

$$\begin{bmatrix} \dfrac{\partial f(x,y)}{\partial x} \\ \dfrac{\partial f(x,y)}{\partial y} \end{bmatrix} = \begin{bmatrix} -2x+2 \\ -2y+4 \end{bmatrix}$$

Recherche d'un extremum :

$$\begin{bmatrix} \dfrac{\partial f(x,y)}{\partial x} \\ \dfrac{\partial f(x,y)}{\partial y} \end{bmatrix} = 0 \Longleftrightarrow (x,y) = (1,2)$$

Calcul de la matrice hessienne :

$$\begin{bmatrix} \dfrac{\partial^2 f(x,y)}{\partial x^2} & \dfrac{\partial^2 f(x,y)}{\partial x \partial y} \\ \dfrac{\partial^2 f(x,y)}{\partial y \partial x} & \dfrac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

La matrice hessienne est définie négative , le point $(1,2)$ est un maximum.

On reconnaît une autre écriture de f et ainsi quelques caractéristiques comme le minimum.

$$f(x,y) = -x^2 + 2x - y^2 + 4y - 7 = -(x-1)^2 - (y-2)^2 - 3$$

Exemple 3

Calcul du gradient :

$$\begin{bmatrix} \dfrac{\partial f(x,y)}{\partial x} \\ \dfrac{\partial f(x,y)}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x \\ -2y \end{bmatrix}$$

Recherche d'un extremum :

$$\begin{bmatrix} \dfrac{\partial f(x,y)}{\partial x} \\ \dfrac{\partial f(x,y)}{\partial y} \end{bmatrix} = 0 \Longleftrightarrow (x,y) = (0,0)$$

Calcul de la matrice hessienne :

$$\begin{bmatrix} \dfrac{\partial^2 f(x,y)}{\partial x^2} & \dfrac{\partial^2 f(x,y)}{\partial x \partial y} \\ \dfrac{\partial^2 f(x,y)}{\partial y \partial x} & \dfrac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

La matrice hessienne n'est ni définie positive ni définie négative.

On obtient un point-selle.

$$\forall x \neq 0 \quad f(x,0) = x^2 > f(0,0)$$

$$\forall y \neq 0 \quad f(0,y) = -y^2 < f(0,0)$$

Application : Modèle linéaire

Construction de la droite des moindres carrés sur une régression linéaire simple

équation de la régression :

$$y_i = ax_i + b + \epsilon_i \quad i = 1, \dots n$$

On cherche a et b minimisant $\sum\limits_{i=1}^{n} \epsilon_i^2$

$\sum\limits_{i=1}^{n} \epsilon_i^2 = \sum\limits_{i=1}^{n} [y_i - (ax_i + b)]^2$

on pose $f(a, b) = \sum\limits_{i=1}^{n} \epsilon_i^2 = \sum\limits_{i=1}^{n} [y_i^2 - 2y_i(ax_i + b) + (ax_i + b)^2]$

Application : Modèle linéaire (suite)

Calcul du gradient :

$$\begin{bmatrix} \dfrac{\partial f(a,b)}{\partial a} \\ \dfrac{\partial f(a,b)}{\partial b} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} [-2y_i x_i + 2x_i(ax_i + b)] \\ \sum_{i=1}^{n} [-2y_i + 2(ax_i + b)] \end{bmatrix}$$

Recherche d'un extremum :

$$\begin{bmatrix} \dfrac{\partial f(a,b)}{\partial a} \\ \dfrac{\partial f(a,b)}{\partial b} \end{bmatrix} = 0 \Longleftrightarrow \begin{cases} -\sum_{i=1}^{n} y_i x_i + \sum_{i=1}^{n} x_i(ax_i + b) = 0 \\ -\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} (ax_i + b) = 0 \end{cases}$$

Application : Modèle linéaire (suite)

$$\begin{cases} -\sum_{i=1}^{n} y_i x_i + a \sum_{i=1}^{n} x_i^2 + b \sum_{i=1} x_i = 0 \\ -n\bar{Y} + na\bar{X} + nb = 0 \end{cases}$$

Calculs intermédiaires :

$$\begin{aligned} \sum_{i=1}^{n} (x_i - \bar{X})^2 &= \sum_{i=1}^{n} [x_i^2 - 2x_i\bar{X} + \bar{X}^2] \\ &= \sum_{i=1}^{n} x_i^2 - 2\bar{X} \sum_{i=1}^{n} x_i + n\bar{X}^2 \\ &= \sum_{i=1}^{n} x_i^2 - n\bar{X}^2 \end{aligned} \tag{1}$$

$$\begin{aligned} \sum_{i=1}^{n} (x_i - \bar{X})(y_i - \bar{Y}) &= \sum_{i=1}^{n} [x_i y_i - x_i\bar{Y} - y_i\bar{X} + \bar{X}\bar{Y}] \\ &= \sum_{i=1}^{n} x_i y_i - n\bar{X}\bar{Y} - n\bar{X}\bar{Y} + n\bar{X}\bar{Y} \\ &= \sum_{i=1}^{n} x_i y_i - n\bar{X}\bar{Y} \end{aligned} \tag{2}$$

Application : Modèle linéaire (suite)

En utilisant les équations (1) et (2), le système devient :

$$\begin{cases} -\sum_{i=1}^{n}(x_i-\bar{X})(y_i-\bar{Y}) - n\bar{X}\bar{Y} + a\sum_{i=1}^{n}(x_i-\bar{X})^2 + an\bar{X}^2 \\ \qquad +nb\bar{X} = 0 \\ b = \bar{Y} - a\bar{X} \end{cases}$$

$$\begin{cases} -\sum_{i=1}^{n}(x_i-\bar{X})(y_i-\bar{Y}) - n\bar{X}\bar{Y} + a\sum_{i=1}^{n}(x_i-\bar{X})^2 + an\bar{X}^2 \\ \qquad +n(\bar{Y}-a\bar{X})\bar{X} = 0 \\ b = \bar{Y} - a\bar{X} \end{cases}$$

$$\begin{cases} a = \dfrac{\sum_{i=1}^{n}(x_i-\bar{X})(y_i-\bar{Y})}{\sum_{i=1}^{n}(x_i-\bar{X})} \\ b = \bar{Y} - a\bar{X} \end{cases}$$