1 Espace vectoriel - Géométrie

L'algèbre matricielle est d'un usage courant en statistique, c'est un moyen d'expression à la fois simple et imagé. Simple parce que le formalisme est "réducteur" et imagé parce qu'il est possible de représenter les objets manipulés.

Qu'est-ce qu'une matrice ?

Quand un échantillon de n individus et p variables est observé, on collecte des données sous la forme de np nombres réels x_{ij}. Chaque x_{ij} est la valeur observée de la j$ème$ variable ($j = 1, 2, \cdots, p$) sur le i$ème$ individu ($i = 1, 2, \cdots, n$).

- Ces valeurs peuvent être rangées dans un tableau rectangulaire à n lignes et p colonnes dans lequel x_{ij} est l'élément à l'intersection de la i$ème$ ligne et de la j$ème$ colonne.

  Ce tableau rectangulaire constitue la <u>*matrice*</u> X des données, on la représente par :

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

- Chaque ligne i de X :

$$X_{i.} = [\, x_{i1} \quad x_{i2} \quad \cdots \quad x_{ip} \,]$$

  peut être considérée comme les <u>*coordonnées*</u> d'un point ou comme un <u>*vecteur*</u> dans un espace de dimension p.

Les n points $i = 1, 2, \cdots, n$ ainsi définis véhiculent la même information que celle qui est contenue dans la matrice des données. Chaque point de l'espace de dimension p représente un *échantillon individuel* et l'espace qui les contient est appelé *espace des individus*.

- Chaque colonne j de X :

$$X_{.j} = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{bmatrix}$$

  détermine les coordonnées d'un point ou un vecteur dans un espace de dimension n.

  Les p points ainsi définis correspondent aux variables, l'espace qui les contient est appelé *espace des variables*. Ces p points véhiculent la même information que celle qui est contenue dans le tableau des données.

- L'espace des individus et l'espace des variables sont dits *espaces duaux*. Ainsi, chaque concept ou raisonnement dans l'un de ces espaces a une image duale dans l'autre. L'une de ces images peut paraître plus simple que l'autre, c'est ce qui justifie les représentations ou les raisonnements faits dans l'un ou l'autre de ces espaces.

Illustration numérique

Soit la matrice :

$$X = \begin{bmatrix} 2 & 4 \\ 4 & 3 \\ 6 & 5 \end{bmatrix}$$

Les lignes de cette matrice représentent des individus (il y en a trois, on les désigne par *ind.1, ind.2, ind.3*), les colonnes représentent des variables (il y en a deux, on les désigne par *var.1, var.2*). Pour chaque individu (chaque ligne), les valeurs prises par *var.1 et var.2* représentent des coordonnées dans un système de deux axes. On peut donc représenter *ind.1, ind.2, ind.3* dans le système d'axes défini par *var.1 et var.2*, c'est l'espace des individus. De même, pour chaque variable (chaque colonne), les valeurs prises par *ind.1, ind.2, ind.3* représentent les coordonnées de *var.1 et var.2* dans le système d'axes défini par les trois individus, c'est l'espace des variables.
Sur la figure 1.1 on donne les représentations de ces données, successivement dans l'espace des individus et dans l'espace des variables.

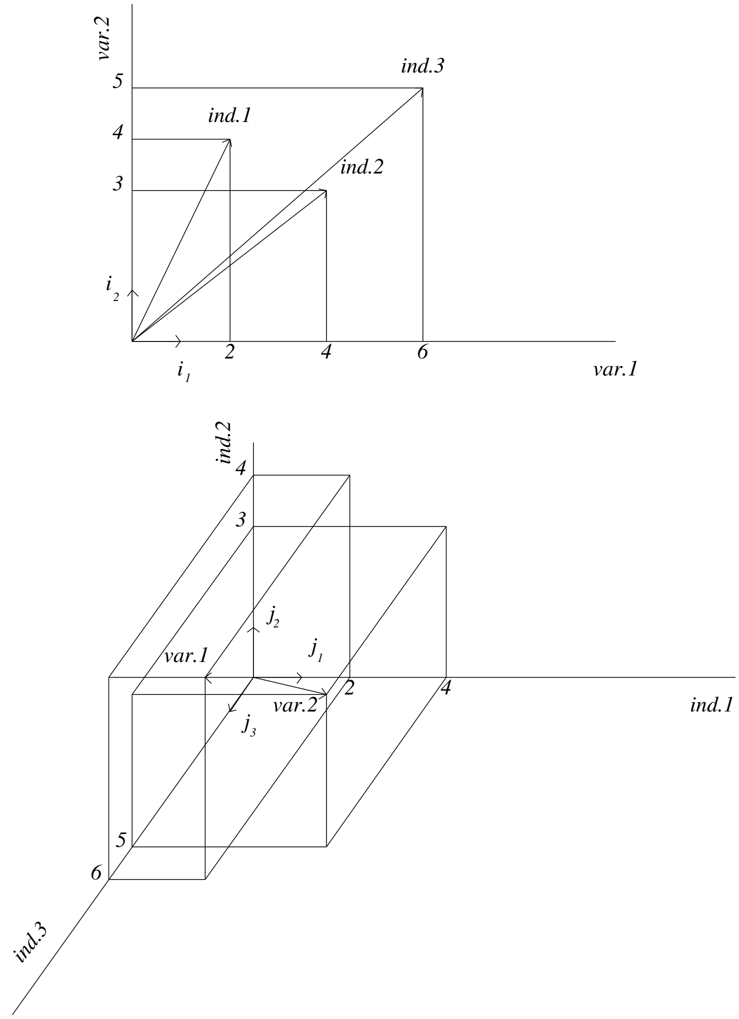

Figure 1.1 : Représentation des individus puis des variables

Les lignes (resp. les colonnes) d'une matrice sont des vecteurs lignes (resp. colonnes), nous allons nous intéresser à eux. Notre objectif est de montrer qu'un vecteur peut être interprété géométriquement. Habituellement, dans le plan de la feuille, on représente un vecteur par une flèche. Deux flèches de même longueur, de même direction et de même sens représentent des vecteurs égaux que l'on désigne par des lettres surmontées de flèches telles que : $\vec{U} = \vec{V}$. Cette notation est d'autant plus lourde que l'on est parfois amené à préciser l'origine et l'extrémité d'un vecteur en écrivant par exemple le vecteur $\vec{V}$ sous la forme $\vec{OA}$ où O est l'origine et A l'extrémité du vecteur $\vec{V}$. Pour éviter cet écueil et pour simplifier l'écriture :

- Les vecteurs seront représentés soit par des lettres majuscules telles que U, V, W, X, $\cdots$, soit par OU, OV, UV, $\cdots$, si on veut préciser l'origine et l'extrémité de ces vecteurs.
- Les *composantes* d'un vecteur seront désignées par des lettres minuscules indicées. Les k composantes du vecteur U seront : $(u_1, u_2, \cdots, u_k)$.

Par ailleurs, les *scalaires* et les *angles* seront désignés par des lettres grecques minuscules telles que α, β, γ, $\cdots$, θ.

1.1 Quelques représentations géométriques

Sur la figure 1.2, sont représentés :

- Dans le plan P de la feuille, un point particulier O appelé *origine* du plan.
- Des segments de droites orientés U, V, $U+V$ d'origine commune O et d'extrémités respectives U, V et $U+V$.

Ces segments de droites ont une *longueur*, une *direction* et un *sens*, ce sont les représentations géométriques des vecteurs U, V, $U+V$.

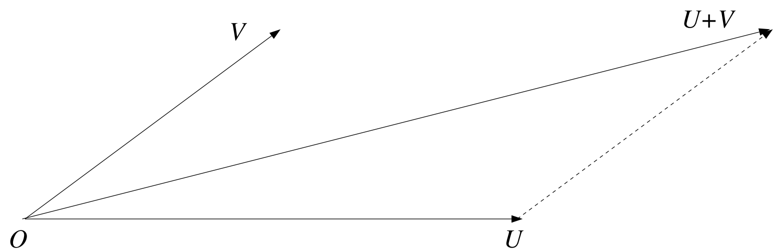

Figure 1.2 : Addition de deux vecteurs.

1.1.1 Egalité de deux vecteurs

- Deux vecteurs sont égaux s'ils ont même longueur, même direction et même sens.

1.1.2 Addition de deux vecteurs

- Géométriquement, l'addition de deux vecteurs U et V est réalisée en traçant à l'extrémité de U (resp. de V) un vecteur égal au vecteur V (resp. de U) et en reliant l'origine O de U (resp. V) à l'extrémité du vecteur placé après U. Sur la figure 1.2 le vecteur somme est noté $U+V$.

1.1.3 Différence de deux vecteurs

- Sur la figure 1.3, sont représentés les vecteurs U, V, $-V$ et $V - U$.

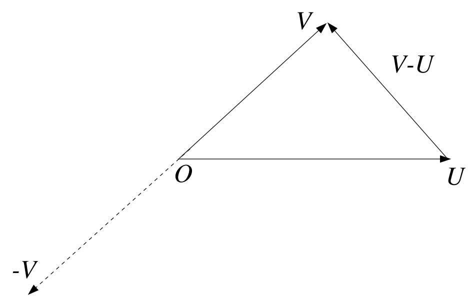

Figure 1.3 : Différence de deux vecteurs.

1.1.4 Multiplication par un scalaire

- Sur la figure 1.4 sont représentés différents cas de multiplications de vecteurs (U, V, W, $\cdots$) et de scalaires (α, β, $\cdots$).

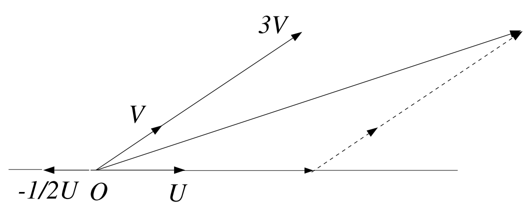

Figure 1.4 : Multiplication par un scalaire.

- Soient U et V deux vecteurs non nuls ayant des directions différentes alors, tout point de P est de la forme $\alpha U + \beta V$.

1.1.5 Longueur, distance, angle, produit scalaire

- C'est à partir des éléments de la figure 1.5 que nous rappelons les notions de longueur, de distance, d'angle et de produit scalaire.

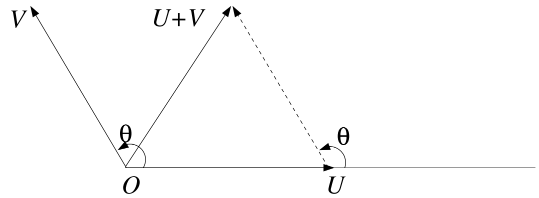

Figure 1.5 : La règle des cosinus.

- On note :

  $|U|$ la <u>*longueur*</u> du vecteur U

  $|V-U|$ la <u>*distance*</u> qui sépare les extrémités U et V des vecteurs OU et OV (cf. figure 1.3).

On considère sur la figure 1.5 le triangle de sommets O, U et $U+V$. La règle des cosinus dit que :

$$|U+V|^2 \ = \ |U|^2 + |V|^2 + 2|U||V|cos(\theta) \qquad (1)$$

où θ est l'angle formé par les vecteurs U et V. Quand l'angle θ est droit, on dit que les vecteurs U et V sont <u>*orthogonaux*</u>. Dans ce cas la formule (1) se réduit à :

$$|U+V|^2 \ = \ |U|^2 + |V|^2 \qquad (2)$$

on reconnaît le théorème de Pythagore.

C'est $2|U||V|cos(\theta)$ qui fait la différence entre les formules (1) et (2). C'est cette quantité qui décrit fondamentalement la relation entre les vecteurs U et V. Cette quantité est (au coefficient 2 près) appelée <u>*produit scalaire*</u> des vecteurs U et V. On la note $< U, V >$:

$$< U, V > = \ |U||V|cos(\theta) \qquad (3)$$

En tenant compte de (3), (1) s'écrit :

$$|U+V|^2 \ = \ |U|^2 + |V|^2 + 2 < U, V > \qquad (4)$$

qui est l'analogue de l'identité remarquable bien connue : $(a+b)^2 \ = \ a^2 + b^2 + 2ab$.

L'intérêt fondamental du produit scalaire apparaît clairement quand on sait que sa connaissance implique celles des longueurs, des distances et des angles entre deux vecteurs.

- *Longueur :* $|U| \ = < U, U >^{1/2}$

- *Distance :* $|V - U| = [< U,U > + < V,V > -2 < U,V >]^{1/2}$
- *Angle :* $cos(\theta) = \dfrac{< U,V >}{(< U,U >< V,V >)^{1/2}}$

Le produit scalaire satisfait aux propriétés suivantes (notées P_1 à P_4) :

- P_1 : $< U,V > = < V,U >$
- P_2 : $< U,U > \geq 0, \quad < U,U > = 0$ si et seulement si $U = 0$
- P_3 : $< \alpha U + \beta V, W > = \alpha < U,W > + \beta < V,W >$
- P_4 : U et V sont orthogonaux si et seulement si : $< U,V > = 0$

1.2 Notion d'espace vectoriel

La géométrie plane peut s'exprimer en termes de vecteurs, d'addition vectorielle, de multiplication d'un vecteur par un scalaire et de produit scalaire ; deux vecteurs non colinéaires déterminent un plan ; $\cdots$. Toutes ces notions sont généralisables à trois dimensions et au-delà de trois dimensions. Par exemple, trois vecteurs non coplanaires engendrent un espace à trois dimensions. La *dimension* d'un espace est en général égale au nombre minimum de vecteurs requis pour déterminer l'espace.

Un espace vectoriel est une abstraction de cette approche géométrique. On dit que l'on a un espace vectoriel quand :

- On a une collection d'éléments, appelés vecteurs, qui peuvent être additionnés et multipliés par des scalaires.

Dans cette approche, on ne décrit pas ce que sont les vecteurs, on décrit comment ils se comportent. Toute chose qui se comporte de cette façon est un vecteur.

Deux concepts complémentaires émergent :

- Bien que d'inspiration géométrique, les éléments d'un espace vectoriel n'ont pas toujours besoin d'apparaître dans une représentation géométrique. On peut représenter un plan sans représenter les vecteurs qui déterminent ce plan.
- Quand un produit scalaire est donné, on définit une *géométrie*. Dans ce cas, on peut interpréter géométriquement les éléments d'un espace vectoriel quelle que soit leur "vraie" nature.

Par exemple, si on considère deux variables aléatoires X et Y, leur somme et leur différence sont aussi des variables aléatoires, il en est de même pour $\alpha X + \beta Y$ où α et β sont des scalaires. Des variables aléatoires peuvent donc être perçues comme des vecteurs et si un produit scalaire approprié est défini, on peut les interpréter géométriquement (des combinaisons linéaires de deux variables aléatoires sont représentées par un plan, $\cdots$).

1.3 Projections

Dans ce paragraphe, nous notons $\hat{U}$, $\hat{V}$, $\hat{W}$, $\cdots$ les projections de U, V, W, $\cdots$ sur un plan P (représenté par un parallélogramme) contenant une origine O (figure 1.6). Ces projections faites parallèlement à une direction D perpendiculaire au plan P sont dites *projections orthogonales*.

Deux propriétés (notées P_1, P_2) définissent les projections orthogonales sur le plan P :

- P_1 : la projection orthogonale $\hat{U}$ de U sur le plan P est dans le plan P.
- P_2 : le vecteur $U - \hat{U}$ est orthogonal à tout vecteur de P. Soit :

$$< V, U - \hat{U} > = \ 0 \text{ pour tout vecteur } V \text{ dans le plan } P.$$

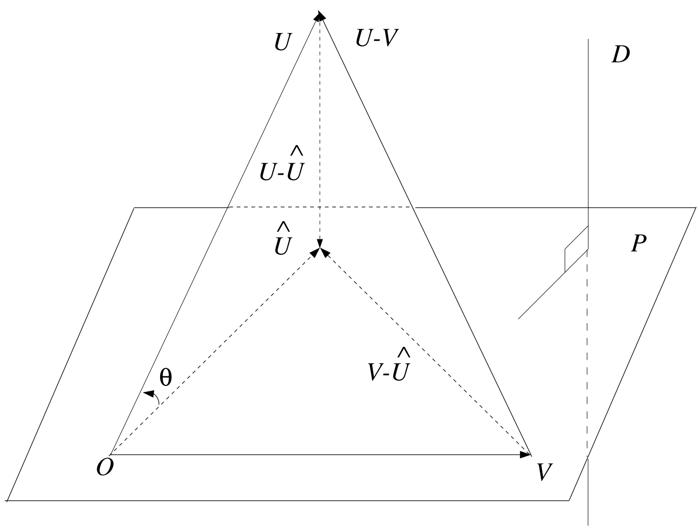

Figure 1.6 : Projections orthogonales.

Les propriétés importantes des projections sont contenues dans les théorèmes suivants (notés T_1 à T_7) :

- T_1 : Le vecteur U est dans le plan P si et seulement si $\hat{U} = U$.
- T_2 : Le vecteur $U \neq 0$ est orthogonal au plan P si et seulement si $\hat{U} = 0$.
- T_3 : La projection sur le plan P du vecteur $\hat{U}$ est égale à $\hat{U}$.
- T_4 : $|U|^2 = |\hat{U}|^2 + |U - \hat{U}|^2$.
- T_5 : Pour tout vecteur U sa projection orthogonale sur le plan P est unique.
- T_6 : Parmi tous les vecteurs du plan P ayant leur origine en O, c'est le vecteur $\hat{U}$ qui est le plus proche de U au sens de la plus petite distance entre l'extrémité U du vecteur OU et l'extrémité $\hat{U}$ du vecteur $O\hat{U}$ contenu dans le plan P.
- T_7 : Les projections sont linéaires : $\alpha \widehat{U + V} = \alpha\hat{U} + \hat{V}$.

1.4 Géométrie analytique

Les notions de vecteurs et de produit scalaire n'ont été utilisées que pour introduire les aspects descriptifs des représentations géométriques. Pour les aspects opératoires, les calculs, il faut traduire ces concepts en termes de *coordonnées*. Nous allons passer brièvement en revue la notion de coordonnées dans un espace à trois dimensions que nous généraliserons ensuite à des espaces à n dimensions.

Coordonnées

Considérons l'espace vectoriel E de dimension 3 défini de la façon suivante :

- 1 - Les vecteurs de E sont définis par les triplets de la forme $U = (u_1, u_2, u_3)$, $V = (v_1, v_2, v_3)$, $\cdots$. Soit :

$$U = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad V = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad \cdots$$

- 2 - L'addition de deux vecteurs est définie composante par composante :

$$(U + V) = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix}$$

- 3 - La multiplication par un scalaire α est définie par :

$$\alpha U = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \alpha u_3 \end{bmatrix}$$

- 4 - Le produit scalaire de deux vecteurs U et V est défini par :

$$< U, V > = u_1 v_1 + u_2 v_2 + u_3 v_3 \qquad (5)$$

On peut vérifier que les définitions d'un espace vectoriel de dimension finie sont satisfaites pour E. La longueur d'un vecteur U est alors donnée par :

$$|U|^2 = u_1^2 + u_2^2 + u_3^2 \qquad (6)$$

Considérons les vecteurs :

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

D'après (5) et (6), nous voyons que :

$$|I_1|^2 = |I_2|^2 = |I_3|^2 = 1$$

$$< I_1, I_2 > = < I_1, I_3 > = < I_2, I_3 > = 0 \qquad (7)$$

Les vecteurs I_1, I_2 et I_3 sont des vecteurs *unitaires* mutuellement orthogonaux. Ainsi, tout vecteur U peut s'écrire comme une combinaison de I_1, I_2 et I_3 :

$$U = u_1 I_1 + u_2 I_2 + u_3 I_3 \qquad (8)$$

On dit que I_1, I_2 et I_3 forment une *base* de l'espace vectoriel. La figure 1.7 montre géométriquement la relation (8). Les axes I_1, I_2 et I_3 sont mutuellement orthogonaux. Les coefficients de (8) sont les *coordonnées* de U pour ce système d'axes.

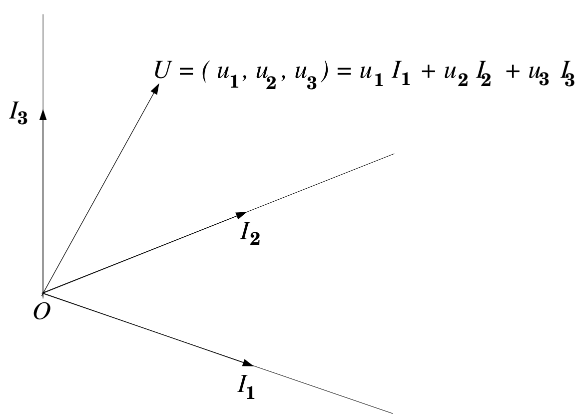

Figure 1.7 : Coordonnées et axes.

Tout triplet de nombres (u_1, u_2, u_3) peut être interprété géométriquement, on peut le considérer comme un vecteur dans un espace déterminé par (8). On peut vérifier que toutes les propriétés qui dérivent du produit scalaire sont en accord avec les concepts de la géométrie ordinaire (géométrie *Euclidienne*). En particulier, les vecteurs U et V sont orthogonaux quand la relation (5) est égale à zéro ; le résultat classique de géométrie analytique :

$$cos^2(\theta) \;=\; \frac{(u_1v_1 + u_2v_2 + u_3v_3)^2}{(u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2)} \qquad (9)$$

est en accord avec celui que suggère (5).

L'équation (6) est la généralisation à trois dimensions du théorème de Pythagore.

Nous avons vu que des vecteurs et un produit scalaire déterminent une géométrie. De ce point de vue, il n'y a pas de raisons de se limiter à trois dimensions. Toute collection de n nombres u_1, u_2, $\cdots$, u_n peut être considérée comme un vecteur

$$U \;=\; \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

- L'addition et la multiplication par un scalaire sont définies sur les coordonnées.
- Les axes ont pour vecteurs unitaires :

$$I_1 \;=\; \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad I_2 \;=\; \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \cdots \quad I_n \;=\; \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

- L'espace vectoriel à n dimensions est constitué par l'ensemble de toutes les combinaisons linéaires de I_1, I_2, $\cdots$, I_n.
- Le produit scalaire des vecteurs U et V est défini par :

$$< U, V > = \; u_1v_1 + u_2v_2 + \cdots + u_nv_n \qquad (10)$$

  On peut calculer des produits scalaires au moyen de la formule (10) si on connaît les coordonnées. Cette formule correspond à un ensemble particulier d'axes. Si nous étions dans un autre système d'axes orthogonaux, chaque point U aurait un nouvel ensemble de coordonnées qu'il faudrait utiliser pour calculer le produit scalaire et donc, les longueurs, les angles.

  Cette formule (10) n'est valable que si on est dans un système d'axes orthogonaux, si tel n'était pas le cas, la formule serait plus compliquée.

Les principes géométriques fondamentaux qui découlent du produit scalaire tels que les longueurs, les angles sont indépendants du système d'axes choisi. Par exemple, l'angle que font deux vecteurs entre eux ne dépend pas du système d'axes de référence.

Le produit scalaire (10) est dit produit scalaire *Euclidien* car, pour $n = 3$, il correspond à la géométrie Euclidienne classique. C'est de loin le produit scalaire le plus utilisé pour des $n - uplets$ de nombres, mais beaucoup d'autres produits scalaires peuvent être définis.

Toute fonction des U, V, $\cdots$, qui satisfait à $< \alpha U + \beta V, W > = \ \alpha < U, W > + \beta < V, W >$ est un produit scalaire convenable.