Table des matières

1. Formes bilinéaires et formes quadratiques

1.1. **Définitions.** Soit E un espace vectoriel sur K ($K = R$ ou C).

Une **forme bilinéaire** sur E est une application $\varphi : E \times E \to K$ linéaire par rapport à chacune des deux variables.

Une forme bilinéaire φ est **symétrique** si $\varphi(x, y) = \varphi(y, x)$ pour tout $x, y \in E$. Si $\varphi(x, y) = -\varphi(y, x)$ pour tous $x, y \in E$, la forme est dite **anti-symétrique** (ou **alternée**).

Chaque forme bilinéaire s'écrit comme la somme d'une forme symétrique et d'une forme anti-symétrique : $\varphi = \varphi_+ + \varphi_-$, où

$$\varphi_+(x,y) = \frac{1}{2}(\varphi(x,y) + \varphi(y,x)) \qquad \varphi_-(x,y) = \frac{1}{2}(\varphi(x,y) - \varphi(y,x)).$$

Pour une forme bilinéaire symétrique on définit la **forme quadratique associée** $q_\varphi : E \to K$: $q_\varphi(x) = \varphi(x,x)$.

La forme bilinéaire symétrique est déterminée par la forme quadratique associée :

$$\varphi(x,y) = \frac{1}{4}[q_\varphi(x+y) - q_\varphi(x-y)]$$

(*"identité de polarisation"*).

L'ensemble de toutes les formes bilinéaires (ou de formes symétriques, ou anti-symétriques, ou quadratiques) est un espace vectoriel sur K : si $\varphi_1, ..., \varphi_k$ sont des formes bilinéaires et $a_1, ...a_k$ des scalaires, $a_1\varphi_1 + ... + a_k\varphi_k$ est une forme bilinéaire.

Exemples. 1. Si f et g sont deux formes linéaires, $\varphi(x,y) = f(x)g(y)$ est une forme bilinéaire.
2. Soit E l'espace des matrices $k \times n$; alors $\varphi(A,B) = tr(AB)$ et $\varphi(A,B) = tr({}^tAB)$ sont des formes bilinéaires symétriques.
3. Soit $E = C([a,b],K)$ l'espace des fonctions continues sur $[a,b]$, soit $p \in C([a,b],K)$. Alors $\varphi(f,g) = \int_a^b f(t)g(t)p(t)dt$ est une forme bilinéaire symétrique.
4. Le déterminant en dimension 2 : $\varphi(x,y) = x_1y_2 - x_2y_1$ - une forme alternée.

1.2. **Expression en coordonnées.** On suppose que dim $E = n < \infty$.
Soit $\mathcal{B} = (e_1, ..., e_n)$ une base de E, $x = \sum_1^n x_ie_i$, $y = \sum_1^n y_ie_i$.
Alors $\varphi(x,y) = \sum_{i,j=1}^n x_iy_j\varphi(e_i,e_j) = \sum_{i,j=1}^n a_{ij}x_iy_j$ où $a_{ij} = \varphi(e_i,e_j)$.
La matrice $A = (a_{ij}) = (\varphi(e_i,e_j))$ est la *matrice de la forme bilinéaire* φ dans la base $\mathcal{B}$.
La forme φ est symétrique si et seulement si sa matrice (dans n'importe quelle base) est symétrique : $a_{ij} = a_{ji}$, ou ${}^tA = A$. Si φ est symétrique, la forme quadratique associée s'écrit : $q_\varphi(x) = \sum_{i,j=1}^n a_{ij}x_ix_j$.
Soit X la colonne des composantes du vecteur x : ${}^tX = (x_1, ..., x_n)$. Alors on peut écrire φ à l'aide de la multiplication matricielle :

$$\varphi(x,y) = {}^tXAY$$

1.3. **Changement de base (changement linéaire de coordonnées).** Soit $\mathcal{B}' = (e'_1, ..., e'_n)$ une autre base de E, soit X' et Y' les colonnes des coordonnées des vecteurs x et y dans la base $\mathcal{B}'$.
On a $X = PX'$ et $Y = PY'$, où P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Alors $\varphi(x,y) = {}^tXAY = {}^tX'{}^tPAPY' = {}^tX'A'Y'$ où

$$A' = {}^tPAP$$

est la matrice de la forme φ dans la base $\mathcal{B}'$.
(Noter que si A est symétrique, A' l'est aussi.)
Plus généralement, on peut effectuer une substitution linéaire $X = CX'$ et $Y = CY'$, où $X' = {}^t(x'_1, ..., x'_k)$ et C est une matrice $n \times k$. Cela donne une forme bilinéaire en k variables $\varphi'(x',y') = {}^tX'A'Y'$ où $A' = {}^tCAC$.
Remarque : Poser $X = CX'$ signifie à remplacer les variables $x_1, ..., x_n$ par des formes linéaires en $x'_1, ..., x'_k$.
Si on définit l'application linéaire $f : K^k \to K^n$ par $f(X') = CX$, alors $\varphi'(x',y') = \varphi(f(x'), f(y'))$.

.

1.4. **Équivalence des formes.** Deux formes bilinéaires φ et φ' définies dans E et E' sont dites **équivalentes** si il existe un isomorphisme $f : E \to E'$ tel que $\varphi(x,y) = \varphi'(f(x), f(y))$.

Si $\dim(E) < \infty$, les formes φ et ψ sont équivalentes si leurs matrices A et B sont liées par $B = {}^tPAP$ avec P inversible (autrement dit, si on peut trouver deux bases dans lequelles φ et ψ ont la même matrice).

1.5. **Rang.** On appelle **rang** d'une forme bilinéaire le rang de sa matrice (il ne dépend pas du choix de la base). On dit que la forme est **non-dégénérée** si son rang est égal à la dimension de E.

Pour une forme φ symétrique son **noyau** est défini par

$\mathrm{Ker}\ \varphi = \{x \in E : \forall y \in E, \varphi(x,y) = 0\}$.

Le noyau de φ est le noyau de (l'application linéaire définie par) la matrice de φ. On a : rang (φ) + dim (Ker φ) = dim (E).

Lemme. *Caractérisation du noyau en termes de la forme quadratique :*

$x \in \mathrm{Ker}\ (\varphi)$ si et seulement si $q_\varphi(x+y) = q_\varphi(y)$ pour tout $y \in E$.

.

1.6. **Vecteurs orthogonaux.** Soit φ une forme bilinéaire symetrique. Les vecteurs x et y sont **orthogonaux** si $\varphi(x,y) = 0$. Une base est dite **orthogonale** si ses vecteurs sont deux à deux orthogonaux.

Dans une base orthogonale la forme s'écrit $\varphi(x,y) = \sum_1^n a_i x_i y_i$, et sa matrice est diagonale. La forme quadratique associée devient alors une combinaison linéaire de carrés : $q(x) = \sum_1^n a_i x_i^2$.

Le rang de φ (ou de q) est le nombre de coefficients a_i non-nuls. Le noyau de φ est engendré par les vecteurs de base e_i pour lesquels $a_i = 0$.

.

1.7. **Orthogonalisation de Gauss (réduction en carrés).** L'orthogonalisation de Gauss permet de fabriquer une base orthogonale pour la forme quadratique $q_\varphi(x) = \sum_{i,j=1}^n a_{ij} x_i x_j$, par des changements de coordonnées successives. On se rappelle que la matrice $A = \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix}$ est symétrique.

Cas 1. Il existe i tel que $a_{ii} \neq 0$. Soit $a_{11} \neq 0$ (quitte à changer la numérotation). On écrit

$$\begin{aligned} q_\varphi(x) &= \sum_{i,j=1}^n a_{ij} x_i x_j \\ &= a_{11} x_1^2 + 2 \sum_{j=2}^n a_{1j} x_1 x_j + \sum_{i,j=2}^n a_{ij} x_i x_j \\ &= a_{11} \left(x_1 + \frac{1}{a_{11}} \sum_{j=2}^n a_{1j} x_j \right)^2 + \sum_{i,j=2}^n a_{ij} x_i x_j - \left(\frac{1}{a_{11}} \sum_{j=2}^n a_{1j} x_j \right)^2 \\ &= a_{11} y_1^2 + q_1(x_2, ..., x_n), \end{aligned}$$

où $y_1 = x_1 + \frac{1}{a_{11}} \sum_{j=2}^n a_{1j} x_j$.

Ensuite il reste à diagonaliser la forme $q_1(x_2, ..., x_n)$ (récurrence). De plus, la famille $(y_1, x_2, \ldots, x_n)$ est libre.

Cas 2. $a_{ii} = 0$ pour tout i . Si $a_{1j} = 0$ pour tout j, la variable x_1 n'apparait pas dans la forme et la récurrence s'applique. Soit $a_{1j} \neq 0$ Pour simplifier, soit $j = 2$.

On écrit

$$\begin{aligned}
q_\varphi(x) &= \sum_{i,j=1}^{n} a_{ij}x_ix_j \\
&= 2a_{12}x_1x_2 + \sum_{j=3}^{n} 2a_{1j}x_1x_j + \sum_{j=3}^{n} 2a_{2j}x_2x_j + \sum_{i,j=3}^{n} a_{ij}x_ix_j \\
&= 2a_{12}\left(x_1 + \frac{1}{a_{12}}\sum_{j=3}^{n} a_{2j}x_j\right)\left(x_2 + \frac{1}{a_{12}}\sum_{j=3}^{n} a_{1j}x_j\right) \\
&\quad + \sum_{i,j=3}^{n} a_{ij}x_ix_j - \frac{2}{a_{12}}\left(\sum_{j=3}^{n} a_{1j}x_j\right)\left(\sum_{j=3}^{n} a_{2j}x_j\right) \\
&= 2a_{12}y_1y_2 + q_2(x_3, ..., x_n).
\end{aligned}$$

où $y_1 = x_1 + \frac{1}{a_{12}}\sum_{j=3}^{n} a_{2j}x_j$ et $y_2 = x_2 + \frac{1}{a_{12}}\sum_{j=3}^{n} a_{1j}x_j$. Ensuite on pose $z_1 = y_1 + y_2$, $z_2 = y_1 - y_2$ et on a $y_1y_2 = \frac{1}{4}(z_1^2 - z_2^2)$.

Après cela il reste à diagonaliser la forme $q_2(x_3, ..., x_n)$. De plus, la famille $(y_1, y_2, x_3, \dots, x_n)$ est libre.

On continue ainsi jusqu'à ce qu'il ne reste que la forme nulle. On obtient $q(x) = \sum_{i=1}^{m} b_iy_i^2$, avec $m \leq n$ et $y_i = \sum_{j=1}^{n} c_{ij}x_j$. Aussi, la famille $(y_1, \dots, y_m, x_{m+1}, \dots, x_n)$ est libre. Posant $y_i = x_i$ et $b_i = 0$ pour $i = m+1, \dots, n$ on a $q(x) = \sum_{i=1}^{n} b_iy_i^2$, et la famille $(y_1, \dots, y_n)$ est libre.

Posons :

$$B = \begin{pmatrix} b_1 & & \\ & b_2 & \\ & & \ddots \end{pmatrix} \qquad C = (c_{ij}) = \begin{pmatrix} c_{11} & c_{12} & \cdots \\ c_{21} & c_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix}$$

Donc A est symétrique, B est diagonale, C est inversible, et $q(x) = {}^tXAX = {}^tYBY$, où $Y = CX$. Autrement dit, $X = C^{-1}Y$, $B = {}^tC^{-1}AC^{-1}$, et C^{-1} est la matrice de passage de la base dans laquelle q est donnée vers une base q-orthogonale.

1.8. **Équivalence des formes quadratiques.** Deux formes quadratiques sont **équivalentes** si les formes bilinéaires symétriques associées sont équivalentes : q_1 et q_2 sont équivalentes si il existe un isomorphisme $f : E \to E$ tel que $q_2(x) = q_1(f(x))$. Les matrices A_1 et A_2 des deux formes sont liées par $A_2 = {}^tPA_1P$ avec P inversible.

Pour une forme réduit en carrés on écrit
$q(x) = \sum_{i=1}^{k} a_ix_i^2$ avec $a_i \neq 0$, $i = 1, ..., k$. Donc k est le rang de q.

.

Équivalence sur C. En posant $\tilde{x}_i = \sqrt{a_i}x_i$ on obtient la forme réduite : $q(x) = \sum_{i=1}^{k} \tilde{x}_i^2$.

Corollaire. Deux formes quadratiques sur C sont équivalentes si et seulement si elles ont le même rang.

.

Formes quadratiques sur R. Signature. On regroupe les coefficients positifs et négatif et on écrit
$q(x) = \sum_{i=1}^{r} a_ix_i^2 - \sum_{i=r+1}^{r+s} a_ix_i^2$ avec $a_i > 0$, $i = 1, ..., k$ et $r + s = k$.

Théorème de Sylvester. Les entiers r et s (le nombre de carrés positif et négatifs) sont indépendants du choix de la base $q-$orthogonale.

Démonstration. Soit $q(x) = \sum_{i=1}^{r} a_i x_i^2 - \sum_{i=r+1}^{k} a_i x_i^2$ et $q(x) = \sum_{i=1}^{r'} b_i y_i^2 - \sum_{i=r'+1}^{k} b_i y_i^2$ deux réductions en carrés, avec $a_i > 0$ et $b_i > 0$ dans les bases $(e_1, ..., e_n)$ et $(e'_1, ..., e'_n)$ respectivement.

Soit $E_+ = \text{Vect}(e_1, ..., e_r)$ et $E_- = \text{Vect}(e'_{r'+1}, ..., e'_n)$. Si $x \in E_+$, on a $q(x) > 0$ sauf si $x = 0$. Si $x \in E_-$, on a $q(x) \leq 0$. Donc $E_+ \cap E_- = \{0\}$. Par conséquent, $dim(E_+) + dim(E_-) \leq dim(E)$, donc $r + n - r' \leq n$, donc $r \leq r'$. Le même raisonnement donne $r' \leq r$, donc $r = r'$. □

En posant $\tilde{x}_i = \sqrt{a_i} x_i$ on obtient la *forme réduite :*

$$q(x) = \sum_{i=1}^{r} \tilde{x}_i^2 - \sum_{i=r+1}^{r+s} \tilde{x}_i^2.$$

Définition. Le couple (r, s) s'appelle **signature** de la forme quadratique.

On a $r + s = \text{rang}(q)$.

Corollaire. *Deux formes quadratiques sur R sont équivalentes si et seulement si elles ont la même signature.*

1.9. **Formes quadratiques positives.** La forme quadratique q est dite **positive** si $q(x) \geq 0$ pour tout $x \in E$ (donc, si $s = 0$).

La forme quadratique q est dite **définie positive** si $q(x) > 0$ pour tout x non-nul (donc, si $r = dim(E)$).

.

Lemme. *Caractérisation "intrinseque" de la signature.* r (respectivement, s) est égal à la dimension maximale d'un sous-espace F tel que la restriction de q (respectivement, de $-q$) sur F soit définie positive.

En termes matriciels, A est positive si ${}^tXAX \geq 0$ pour tout X ; A est définie positive si ${}^tXAX > 0$ pour tout $X \neq 0$.

Remarque : pour toute matrice C la matrice $A = {}^tCC$ est positive ; ${}^tCC$ est définie positive si et seulement si C est inversible.

1.10. **Orthogonalisation de Gauss pour les formes définie positives.** Si $q_\varphi(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ est définie positive, on a $a_{ii} > 0$ pour tout i. Donc dans l'algorithme de Gauss on rencontre uniquement le cas 1 (voir **1.7.**). La matrice de changement de variables est à chaque étape triangulaire (supérieure) ; la matrice de passage P vers la base orthonormée dans laquelle q est la somme des carrés est donc triangulaire supérieure : ${}^tPAP = I_n$. Soit $C = P^{-1}$. On a $A = {}^tCC$.

Théorème de factorisation triangulaire (Gauss-Cholesky).

Pour toute matrice A symétrique définie positive il existe une unique matrice C triangulaire supérieure à diagonale positive telle que $A = {}^tCC$.

1.11. **Exemple : étude des extréma.** Soit $f(x_1, ..., x_n)$ une fonction de classe C^2 dans R^n. Soit $0 = (0, ..., 0)$ un point critique : $\frac{\partial f}{\partial x_i}(0) = 0$, $i = 1, ..., n$. On considère le développement limité de f en 0 à l'ordre 2 :

$f(x) = f(0) + \frac{1}{2} \sum_{i,j} h_{i,j} x_i x_j + o(\|x\|^2)$, où $h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(0)$.

La forme quadratique $H(x) = \sum_{i,j} h_{i,j} x_i x_j$ est *la forme Hessienne* de f en 0.

Proposition. (i) Si f admet un minimum local en O, H admet un minimum en 0 et donc H est positive.

(ii) Si H admet un minimum strict en 0 et donc est définie positive, f admet un minimum local strict en 0.

2. Produit scalaire et espaces euclidiens

2.1. Définitions.

Définition. Soit E un R-espace vectoriel. Un **produit scalaire** dans E est une forme bilinéaire symétrique définie positive, noté $\langle ., . \rangle$.

La **norme** associée est définie par $\|x\|^2 = \sqrt{\langle x, x \rangle}$ - i.e., la racine carrée de la forme quadratique associée.

On observe que si $\langle \cdot, \cdot \rangle$ est un produit scalaire alors

$$x = 0 \iff \forall y \ \langle x, y \rangle = 0 \iff \|x\| = 0.$$

Théorème (Inégalité de Cauchy-Schwartz)**.** *Soit $\langle ., . \rangle$ un produit scalaire, $\|.\|$ la norme associée. Alors $|\langle x, y \rangle| \leq \|x\|\|y\|$, et on a égalité si et seulement si x et y sont colinéaires.*

Démonstration. Si $x = 0$: pas de problème. On suppose donc que $x \neq 0$, et on étudie $\langle tx + y, tx + y \rangle$. □

Corollaire. *Soit $\langle ., . \rangle$ un produit scalaire, $\|.\|$ la norme associée. Alors $\|.\|$ vérifie l'inégalité du triangle, et c'est bien une norme. Elle induit donc une* **distance euclidienne** *dans E, définie par $d(x, y) = \|x - y\|$*

Le produit scalaire est déterminé par la norme associée (identité de polarisation) :

$$\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$

Remarque : une norme dans E n'est pas en général associée à un produit scalaire. Elle l'est si et seulement si elle vérifies l'identité du parallélogramme : $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$.

Exemples : 1. Produit scalaire canonique dans R^n : $\langle x, y \rangle = \sum_1^n x_i y_i$, $\|x\| = \sqrt{\sum_1^n x_i^2}$.

2. $E = C([a, b], R)$, $\langle f, g \rangle = \int_a^b f(t)g(t)dt$.

Définition. Un R-espace vectoriel de dimension finie muni d'un produit scalaire s'appelle **espace euclidien**.

2.2. Vecteurs orthogonaux.

Deux vecteurs x et y sont **orthogonaux** si $\langle x, y \rangle = 0$.

Sous-espace orthogonale. Soit $A \subset E$; **l'orthogonal** de A est l'ensemble de vecteurs de E orthogonaux à tous les vecteurs de A :

$A^\perp = \{x \in E : \forall y \in A \text{ on a } \langle x, y \rangle = 0\}$.

Il est claire que $A^\perp$ est un sous-espace vectoriel de E.

Deux sous-espaces E_1 et E_2 sont **orthogonaux** si tout vecteur de E_1 est orthogonal à tout vecteur de E_2. Ceci est équivalent à dire que $E_2 \subset E_1^\perp$ ou que $E_1 \subset E_2^\perp$. Il est évident dans ce cas que $E_1 \cap E_2 = \{0\}$.

.

Définition. Une famille de vecteurs de E est dite **orthogonale** si les vecteurs de cette famille sont deux à deux orthogonaux.

Une famille de vecteurs de E est dite **orthonormée** si elle est orthogonales et tous ses vecteurs sont de norme 1.

Lemme ("Théorème de Pythagore")**.** *Soit $(v_1, \ldots, v_k) \subseteq E$ une famille orthogonale, et $u = \sum_1^k v_i$. Alors $\|u\|^2 = \sum_1^k \|v_i\|^2$.*

Lemme. *Une famille orthogonale sans vecteurs nuls est libre.*

D'une famille orthogonale $(e_1, ..., e_n, ...)$ on peut facilement passer à une famille orthonormée en normalisant les vecteurs e_i : $e'_i = \frac{e_i}{\|e_i\|}$.

Exemple. 1. Dans l'espace des fonctions continues $C([0, 2\pi], R)$ avec le produit scalaire $\langle f, g\rangle = \int_0^{2\pi} f(t)g(t)dt$ la famille $(1, \cos nx, \sin nx)_{n\geq 1}$ est orthogonale et la famille $(\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \sin nx)_{n\geq 1}$ est orthonormée.

2. Dans l'espace des fonctions continues $C([-1, 1], R)$ avec le produit scalaire $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)dt$ la famille des polynômes

$(p_n(t) = \frac{d^n}{dt^n}(t^2 - 1)^n)_{n\geq 0}$ est orthogonale.

2.3. **Coordonnées dans une base orthonormée.** Soit $(e_1, ..., e_n)$ une base **orthonormée** , soit $x = \sum_1^n x_i e_i$, $y = \sum_1^n y_i e_i$. Alors $\langle x, y\rangle = \sum_1^n x_i y_i$, $\|x\|^2 = \sum_1^n \|x_i e_i\|^2 = \sum_1^n x_i^2$ ("théorème de Pythagore") et $x_i = \langle x, e_i\rangle$.

Coordonnées dans une base **orthogonale :** $\langle x, y\rangle = \sum_1^n \langle e_i, e_i\rangle x_i y_i$

$\|x\|^2 = \sum_1^n \langle e_i, e_i\rangle x_i^2$ et $x_i = \frac{\langle x, e_i\rangle}{\langle e_i, e_i\rangle}$.

2.4. **Orthogonalisation de Gram-Schmidt.** Soit $(v_1, ..., v_n, ...)$ une famille libre dans E. On peut construire une famille orthonormée $e_1, ..., e_n, ...$ telle que $Vect(v_1, ..., v_k) = Vect(e_1, ..., e_k)$ pour tout $k \geq 1$. (Autrement dit, e_k est une combinaison linéaire de $v_1, ..., v_k$.)

Construction par récurrence :

On pose $e_1 = \frac{v_1}{\|v_i\|}$; $\tilde{e}_{k+1} = v_{k+1} - \sum_1^k \langle v_{k+1}, e_i\rangle e_i$ et $e_{k+1} = \frac{\tilde{e}_{k+1}}{\|\tilde{e}_{k+1}\|}$.

Corollaire. *Tout espace euclidien admet une base orthonormée. Toute famille orthonormée dans un espace euclidien peut être complétée en une base orthonormée.*

Remarque : le corollaire est faux en dimension infinie : l'espace des fonctions continues $C([a, b], R)$ n'admet pas de base orthogonale au sens algébrique.

2.5. **Projection orthogonale.**

Définition. Soit E muni d'un produit scalaire, $F \subseteq E$ un sous-espace. On appelle une **projection orthogonale** sur F une application $P_F \colon E \to F$ telle que $x - P_F(x) \perp F$ pour chaque $x \in E$.

On observe que $x - P_F(x) \perp F$ ssi $\langle x - P_F(x), z\rangle = 0\rangle$ pour tout $z \in F$, ssi $\langle x, z\rangle = \langle P_F(x), z\rangle$ pour tout $z \in F$.

Proposition. *Soit E muni d'un produit scalaire, $F \subseteq E$ un sous-espace.*

(i) *Si une projection orthogonale $P_F \colon E \to F$ existe alors elle est unique. Dans ce cas l'unique projection orthogonale est linéaire, elle est l'identité sur F, et $\ker P_F = F^\perp$. Autrement dit, $E = F \oplus F^\perp$ et P_F est la projection sur F parallèlement à $F^\perp$.*

(ii) *Si F est de dimension finie alors P_F existe.*

Démonstration. Unicité : si $y, y' \in F$ sont tels que $x - y \perp F$ et $x - y' \perp F$ on obtient $y - y' \perp F$. En particulier $y - y' \perp y - y'$ d'où $\|y - y'\| = 0$ et $y = y'$.

Linéarité : $P_F(\lambda x + y) = \lambda P_F(x) + P_F(Y)$ par unicité, puisque pour tout $z \in F$ on a

$$\begin{aligned}\langle \lambda x + y, z\rangle &= \lambda\langle x, z\rangle + \langle y, z\rangle = \lambda\langle P_F(x), z\rangle + \langle P_F(y), z\rangle \\ &= \langle \lambda P_F(x) + P_F(y), z\rangle.\end{aligned}$$

Si $x \in F$ alors $P_F(x) = x$ par unicité. Finalement, par l'unicité on a $x \in F^\perp$ ssi $x - 0 \perp F$ ssi $P_F(x) = 0$.

Existence : On suppose maintenant que F est de dimension finie. Soit $(e_1, ..., e_n)$ une base orthonormée de F, et définissons $P_F(x) = \sum\langle x_i, e_i\rangle e_i$. On vérifie que $x - P_F(x) \perp e_j$ pour chaque $j = 1, \ldots, n$, d'où $x - P_F(x) \perp F$. $\square$

Remarque. Supposons que P_F existe et soit $x \in E$. Alors $x - P_F(x) \in F^\perp$ et $x - (x - P_F(x)) = P_F(x) \perp F^\perp$. Donc $P_{F^\perp}$ existe et est égale à $\mathrm{Id} - P_F$.

Il en suit en particulier que $E = F^\perp \oplus (F^\perp)^\perp$. Puisque $F \subseteq (F^\perp)^\perp$, on obtient $F = (F^\perp)^\perp$.

Par contre : il est possible en dimension infinie que P_F n'existe pas, et que $(F^\perp)^\perp \supsetneq F$.

Projection orthogonale dans une base quelconque. Le vecteur $y = P_F(x)$ est caractérisé par les conditions $y \in F$ et $\langle y, z\rangle = \langle x, z\rangle$ pour tout vecteur z de F.

Soit $(e_1, ..., e_n)$ une base de F. La dernière condition est donc équivalente à $\langle y, e_j\rangle = \langle x, e_j\rangle$, $j = 1, ..., n$.

Posons $P_F(x) = \sum_1^n y_i e_i$. pour déterminer les coefficients y_i on doit résoudre le système :

$$\sum_1^n y_i \langle e_i, e_j\rangle = \langle x, e_j\rangle, \qquad j = 1, ..., n.$$

La matrice de ce système $G = (\langle e_i, e_j\rangle)$ s'appelle *matrice de Gram.* Soit $(\tilde{e}_1, ..., \tilde{e}_n)$ une base orthonormée et $e_j = \sum_i a_{ij}\tilde{e}_i$, i.e., $a_{ij} = \langle \tilde{e}_i, e_j\rangle$. Soit $A = (a_{ij})$, alors $G = {}^tAA$. En particulier, G est définie positive et $\det G = (\det A)^2$.

2.6. **Projection othogonale et meilleur approximation en moyenne quadratique. Distance à un sous-espace.**

Lemme. *Soit F est un sous-espace de dimension finie et $x \in E$. Le point $y = P_F(x)$ est le point de F le plus proche de x. En d'autres termes, c'est un minimum global strict de la fonction $z \mapsto \|x - z\|$ sur F.*

En particulier, $d(x, F) = d(x, P_F(x))$.

Démonstration. En effet, soit $z \in F$, $z \neq y$. Alors $y - z \in F$ d'où $x - y \perp y - z$ et $\|x - z\|^2 = \|x - y\|^2 + \|y - z\|^2 > \|x - y\|^2$. $\square$

Exemple. Ajustement affine.

Soit $x_1 < x_2 < ... < x_n$ et $S = (x_1, ..., x_n)$.

Soit E l'espace des fonctions définies sur S à valeurs réelles.

Le produit scalaire dans E est défini par $\langle f, g\rangle = \sum_1^n f(x_i)g(x_i)$;

Etant donné f, *l'ajustement affine par les moindres carrés* consite à déterminer une fonction affine $\phi(x) = ax + b$ telle que l'écart $\|f - \phi\|^2 = \sum_1^n [f(x_i) - \phi(x_i)]^2$ soit minimal.

La réponse est donnée par la projection orthogonal sur le sous-espace des fonctions affines. Les coefficients a et b sont les solutions du système linéaire : $\langle \phi, 1\rangle = \langle f, 1\rangle$, $\langle \phi, x\rangle = \langle f, x\rangle$. Plus explicitement,

$na + (\sum x_i)b = \sum f(x_i)$,

$(\sum x_i)a + (\sum x_i^2)b = \sum x_i f(x_i)$.

Exemple. Meilleure approximation en moyenne quadratique par des polynômes trigonométriques.

Un polynôme trigonométrique de dégré $\leq n$ est la somme $p(t) = a_0 + \sum_{i=1}^n (a_k \cos kt + b_k \sin kt)$. Soit $f \in C([0, 2\pi])$ une fonction continue. On cherche on polynôme trigonométrique p de degré $\leq n$ tel que l'écart $\int_0^{2\pi} (f(t) - p(t))^2 dt$ soit minimal.

La réponse est donnée par la projection orthogonal dans $C([0, 2\pi])$ sur le sous-espace des polynômes trigonométriques de dégré $\leq n$; le produit scalaire est $\langle f, g\rangle = \int_0^{2\pi} f(t)g(t)dt$. On a la famille orthogonale : $(1, \cos nx, \sin nx)_{n\geq 1}$. On en déduit les coefficients du polynôme $p(t)$ de meilleure approximation : $a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(t)dt$, $a_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos(kt)dt$, $b_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin(kt)dt$.

2.7. **Inégalité de Bessel et égalité de Bessel-Parseval.**

Soit E un espace euclidien, soit $(e_1, ..., e_n)$ une famille orthonormée.

Soit $x \in E$ et $x_i = \langle x, e_i\rangle$. Il est claire que $\sum_{i=1}^n x_i^2 \leq \|x\|^2$.

Lemme. *Soit $x \in E$. Les propriétés suivantes sont équivalentes :*

(i) $\sum_{i=1}^n x_i^2 = \|x\|^2$.

(ii) $x = \sum_{i=1}^n x_i e_i$.

(iii) *x appartient à l'espace vectoriel engendré par $(e_1, ..., e_n)$ (donc x est une cobinaison linéare des vecteurs $e_1, ..., e_n$).*

Soit maintenant $\dim(E) = \infty$.

Théorème. *Soit $(e_1, ..., e_n, ...)$ une famille orthonormée infinie, $x \in E$ et $x_i = \langle x, e_i \rangle$.*
A. *Pour tout n on a $\sum_{i=1}^n x_i^2 \leq \|x\|^2$ et la série $\sum_{i=1}^\infty x_i^2$ converge.*
B. *Les propriétés suivantes sont équivalentes :*

(i) $\sum_{i=1}^\infty x_i^2 = \|x\|^2$.

(ii) $x = \lim_{n\to\infty} \sum_{i=1}^n x_i e_i$.

(iii) *x appartient à l'adhérence de l'espace vectoriel engendré par la suite $(e_1, ..., e_n, ...)$.*

Démonstration. Posons $y_n = \sum_{i=1}^n x_i e_i$ et $z_n = x - y_n$. Alors $y_n = P_{\mathrm{Vect}(e_1,...,e_n)}(x)$, d'où $y_n \perp z_n$ et

$$\|x\|^2 = \|y_n\|^2 + \|z_n\|^2 \geq \|y\|^2 = \sum_{i=1}^n x_i^2,$$

d'où le premier énoncé. Maintenant :

- **(i) $\Longrightarrow$ (ii):** Si $\sum_{i=1}^\infty x_i^2 = \|x\|^2$ c'est que $\|x\|^2 - \sum_{i=1}^n x_i^2 \overset{n\to\infty}{\to} 0$. Or $\|x\|^2 - \sum_{i=1}^n x_i^2 = \|x - y_n\|^2$, d'où $x = \lim y_n$.
- **(ii) $\Longrightarrow$ (iii):** Clair.
- **(iii) $\Longrightarrow$ (i):** On suppose que pour chaque $\varepsilon > 0$ il existe $w \in \mathrm{Vect}(e_1, \ldots, e_n, \ldots)$ tel que $\|x - w\| < \sqrt{\varepsilon}$. Fixons ε, et exprimons $w = \sum_{i=1}^n w_i e_i$ (où n dépend de w!) Alors $\varepsilon \geq \|x - w\|^2 = \|y_n - w + z_n\|^2 = \|y_n - w\|^2 + \|z_n\|^2 \geq \|z_n\|^2 = \|x\|^2 - \|y_n\|^2 \geq 0$. On a donc montré que pour chaque $\varepsilon > 0$ il existe n tel que $\varepsilon > \|z_n\|$. Puisque la suite $\|y_n\|^2 = \sum_{i=1}^n x_i^2$ est croissante, on obtient $\|y_n\|^2 \to \|x\|^2$, i.e., $\sum_{i=1}^\infty x_i^2 = \|x\|^2$.

□

Egalité de Bessel-Parseval dans une "base" orthogonale.
Si $(e_1, ..., e_n, ...)$ est une famille orthogonale et x appartient à l'adhérence de l'espace vectoriel engendré par la suite $(e_1, ..., e_n, ...)$, alors
$\|x\|^2 = \sum_{i=1}^\infty \frac{\langle x, e_i \rangle^2}{\langle e_i, e_i \rangle}$.

Exemple : séries de Fourier. Avec le produit scalaire $\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)dt$ dans $C([0, 2\pi])$ la famille $(1, \cos nt, \sin nt)_{n\geq 1}$ est orthogonale ; les combinaisons linéaires des ses fonctions - les polynômes trigonométriques - sont denses dans $C([0, 2\pi])$. Soit $a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(t)dt$, $a_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(nt)dt$, $b_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(nt)dt$. On a l'égalité de Parseval :
$\frac{1}{\pi} \int_0^{2\pi} f(t)^2 dt = 2a_0^2 + \sum_{n=1}^\infty (a_n^2 + b_n^2)$.

3. Endomorphismes d'espaces euclidiens

3.1. **L'endomorphisme adjoint.** Soit E un espace euclidien, $\mathrm{End}(E)$ l'espace d'endomorphismes de E (c'est-à-dire, des applications linéaires $E \to E$).

Pour $f \in \mathrm{End}(E)$, on s'intéresse aux formes bilinéaires $\langle x, fy \rangle$ et $\langle fx, y \rangle$.

Définition. Soit $f \in \mathrm{End}(E)$. On appelle **adjoint** de f un endomorphisme $f^* \in \mathrm{End}(E)$ tel que pour tout $x, y \in E$:

$$\langle fx, y \rangle = \langle x, f^* y \rangle.$$

Proposition. (i) *Pour tout $f \in \mathrm{End}(E)$, il existe un unique endomorphisme adjoint f^*.*

(ii) *Si $\mathcal{B}$ est une base orthonormée, on a $M_{\mathcal{B}}(f^*) = {}^t M_{\mathcal{B}}(f)$.*

Démonstration. Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée et $A = (a_{ij}) = M_{\mathcal{B}}(f)$. On a $fe_j = \sum_i a_{ij} e_i$ d'où $\langle e_i, fe_j\rangle = a_{ij}$. Montrons d'abord l'unicité, en supposant que f^* existe.

$$f^* e_i = \sum_{j=1}^{n} \langle f^* e_i, e_j\rangle e_j = \sum_{j=1}^{n} \langle e_i, fe_j\rangle e_j = \sum_{j=1}^{n} a_{ij} e_j$$

d'où $M_{\mathcal{B}}(f^*) = {}^t A$ et l'unicité de f^*.

Pour l'existence, on choisit f^* comme l'unique endomorphisme tel que $M_{\mathcal{B}}(f^*) = {}^t A$.

$$\langle e_i, fe_j\rangle = a_{ij} = \langle f^* e_i, e_j\rangle.$$

Il en découle que $\langle x, fy\rangle = \langle f^* x, y\rangle$ pour tout x, y (et ceci ne dépend plus de la base $\mathcal{B}$!) □

Définition. Un endomorphisme f est dit **auto-adjoint** ou **symétrique** si $f = f^*$.

Un endomorphisme f est symétrique ssi sa matrice dans une (toute) base orthonormée est symétrique.

Exemple. Une projection orthogonale est symétrique.

Lemme. (i) *L'application $f \to f^*$ est linéaire :*

$$(f+g)^* = f^* + g^*, \qquad (\alpha f)^* = \alpha f^*.$$

(ii) *On a*

$$(fg)^* = g^* f^*, \qquad (f^*)^* = f.$$

(iii) *Si f est inversible, f^* l'est aussi et*

$$(f^*)^{-1} = (f^{-1})^*.$$

3.2. **Endomorphismes orthogonaux, matrices orthogonales.**

Définition. Un endomorphisme $f \in \mathrm{End}(E)$ est dit **orthogonal** si $f^{-1} = f^*$.

Une matrice $O \in M_n(\mathbf{R})$ est dites **orthogonale** si $O^{-1} = {}^t O$.

Un endomorphisme f est orthogonal ssi sa matrice dans une (toute) base orthonormée est orthogonale.

Proposition. *Si f, g sont des endomorphismes orthogonaux alors f^{-1} et fg le sont aussi. En particulier, l'ensemble des endomorphismes orthogonaux, noté $O(E)$, est un groupe.*

Pareil pour les matrices. Le groupe des matrices orthogonales est $O_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) : {}^t AA = \mathrm{Id}\}$.

Proposition. *Pour $f \in \mathrm{End}(E)$, les propriétés suivantes sont équivalentes.*

(i) *f est orthogonal.*

(ii) *f est isométrique : $\|fx\| = \|x\|$ pour tout x.*

(iii) *f préserve le produit scalaire : $\langle fx, fy\rangle = \langle x, y\rangle$ pour tout x, y.*

(iv) *f transforme toute base orthonormée en base orthonormée.*

(v) *f transforme une base orthonormée en base orthonormée.*

Démonstration. (i) $\Longrightarrow$ (ii). $\|fx\|^2 = \langle fx, fx\rangle = \langle x, f^* fx\rangle = \langle x, f^{-1} fx\rangle = \langle x, x\rangle = \|x\|^2$.

(ii) $\Longrightarrow$ (iii). Puisque le produit scalaire est détermine par la norme (identité de polarisation).

(iii) $\Longrightarrow$ (iv). Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée. Alors $\langle fe_i, fe_j\rangle = \langle e_i, e_j\rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ (le delta de Kronecker). Donc, $(fe_1, \ldots, fe_n)$ est aussi une famille orthonormée. Elle est donc libre, de taille $n = \dim E$, c'est donc une base.

(iv) $\Longrightarrow$ (v). Clair.

(v) $\Longrightarrow$ (i). Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée telle que $(fe_1, \ldots, fe_n)$ est aussi une base orthonormée, et montrons que $f^*fx = x$ pour tout x. En effet on a $x = \sum_{i=1}^n \langle x, e_i\rangle e_i$ d'où $fx = \sum_{i=1}^n \langle x, e_i\rangle fe_i$. Puisque $(fe_1, \ldots, fe_n)$ est une base orthonormée : $fx = \sum_{i=1}^n \langle fx, fe_i\rangle fe_i = \sum_{i=1}^n \langle f^*fx, e_i\rangle fe_i$. Il en découle que $\langle x, e_i\rangle = \langle f^*fx, e_i\rangle$ pour tout $i = 1, \ldots, n$. Par conséquent :

$$x = \sum_{i=1}^n \langle x, e_i\rangle e_i = \sum_{i=1}^n \langle f^*fx, e_i\rangle e_i = f^*fx.$$

Donc $f^* = f^{-1}$. □

Lemme. *Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée, $\mathcal{B}' = (e'_1, \ldots, e'_n)$ une base quelconque, P la matrice de passage. Alors P est orthogonale ssi $\mathcal{B}'$ est orthonormée.*

Démonstration. Soit f l'endomorphisme qui envoie $\mathcal{B}$ à $\mathcal{B}'$, de sorte que $P = M_{\mathcal{B}}(f)$.

Alors P est une matrice orthogonale ssi f est un endomorphisme orthogonal ssi $\mathcal{B}'$ est une base orthonormée. □

Lemme. *Une matrice $A \in M_n(\mathbf{R})$ est orthogonale ssi ses colonnes (lignes) forment une base orthonormée de $\mathbf{R}^n$ par rapport au produit scalaire standard.*

3.3. Théorème spectral : diagonalisation des endomorphismes et matrices symétriques.

Lemme. (i) *L'orthogonal d'un sous-espace stable par f est stable par f^*. C'est à dire que si $F \subseteq E$ est un s.e.v. tel que $f(F) \subseteq F$ alors $f^*(F^\perp) \subseteq F^\perp$.*

(ii) $\ker f^* = (\operatorname{im} f)^\perp$ *et* $\operatorname{im} f^* = (\ker f^*)^\perp$.

Corollaire. *Si f est symétrique, alors $\ker f = (\operatorname{im} f)^\perp$, et $E = \ker f \oplus \operatorname{im} f$. De plus, l'orthogonal d'un sous-espace stable par f est stable par f, i.e., si $f(F) \subseteq F$ alors $f(F^\perp) \subseteq F^\perp$.*

Lemme. *Un endomorphisme $f \in \operatorname{End}(E)$ est une projection orthogonale ssi $f = f^* = f^2$. Autrement dit, f est une projection orthogonale ssi f est auto-adjoint et* **idempotent** *($f = f^2$).*

Dans ce cas : $f = P_F$ où $F = \operatorname{im} f$.

Démonstration. $\Longrightarrow$: Soit $F \subseteq E$ un s.e.v. Alors on sait déjà que $P_F^2 = P_F$, et $\langle P_F x, y\rangle = \langle P_F x, P_F y\rangle = \langle X, P_F y\rangle$ implique $P_F = P_F^*$.

$\Longleftarrow$: soit $f \in \operatorname{End}(E)$ tel que $f = f^* = f^2$, et posons $F = \operatorname{im} f$. Puisque $f = f^*$ on a $\ker f \perp \operatorname{im} f$, i.e., $\ker f \perp F$. Puisque $f = f^2$ on a $f(x - fx) = 0$, d'où $x - fx \in \ker f$. On obtient $fx \in F$ et $x - fx \perp F$ pour tout x, d'où $f = P_F$. □

Lemme. *Soit f un endomorphisme auto-adjoint. Alors f admet une valeur propre réelle.*

Remarque. Soit E un espace euclidien, $G \subseteq F \subseteq E$ des s.e.v. On a *a priori* deux projections orthogonales sur G, de E et de F, respectivement, que l'on notera : $P_G \colon E \to G$ et $P_G^F \colon F \to G$, ce qui risque d'engendrer la confusion.

Or : si on vérifie les définitions, on voit que $P_G^F = P_G{\restriction_F}$ est la restriction de P_G^E à F. On peut donc se permettre d'écrire $P_G x$ sans risque d'ambiguïté. De surcroît, puisque P_F est l'identité sur G on a $P_F P_G = P_G$, d'où $P_G P_F = P_G^* P_F^* = (P_F P_G)^* = P_G^* = P_G$.

Théorème (Théorème spectral, 1ère formulation)**.** *Soit f un endomorphisme auto-adjoint. Soit $(\lambda_1, \dots, \lambda_k)$ les valeurs propres de f, et soit $E_i = \ker(f - \lambda_i \operatorname{Id})$ le sous-espace propre correspondant à λ_i. Alors*

(i) *Les sous-espaces propres de f sont deux à deux orthogonaux, i.e., $E_i \perp E_j$ pour $i \neq j$. En plus, $E_1 \oplus E_2 \oplus \ldots \oplus E_k = E$.*

(ii) **Décomposition spectrale :** $f = \sum_{i=1}^{k} \lambda_i P_{E_i}$.

Corollaire (Théorème spectral, 2ème formulation)**.** *Soit f un endomorphisme auto-adjoint. Alors f est diagonalisable dans une base orthonormée. Autrement dit, il existe une base orthonormée qui consiste en des vecteurs propres de f.*

Corollaire. *Soit $A \in M_n(\mathbf{R})$ une matrice symétrique. Alors*

(i) *Toutes les valeurs propres de A sont réelles.*

(ii) *A est diagonalisable par une matrice orthogonale : il existe une matrice orthogonale O telle que $O^{-1}AO = {}^tOAO$ est diagonale.*

On observe que la réciproque est vrai également : si O est orthogonale et ${}^tOAO$ est diagonale, A est symétrique.

3.4. **Diagonalisation simultanée d'endomorphismes ou matrices.**

Lemme. *Soit $f, g \in \operatorname{End}(E)$ auto-adjoint, et supposons qu'ils commutent : $fg = gf$. Soit $F \subseteq E$ un espace propre de f. Alors*

(i) *Les s.e.v. F et $F^\perp$ sont stables par g : $g(F) \subseteq F$ et $g(F^\perp) \subseteq F^\perp$.*

(ii) *$P_F g = g P_F$.*

(iii) *Si G est un espace propre de g alors on a $P_F P_G = P_G P_F = P_{F \cap G}$.*

Démonstration. Soit λ la valeur propre telle que $F = \ker(f - \lambda \operatorname{Id})$. Si $x \in F$ alors $fgx = gfx = g(\lambda x) = \lambda gx$ d'où $gx \in F$ également, donc $g(F) \subseteq F$. Puisque g est auto-adjoint, ceci implique que $g(F^\perp) \subseteq F^\perp$.

Pour montrer que $P_F gx = gP_F x$, il faut montrer que $gP_F x \in F$ et $gx - gP_F x \perp F$. On sait déjà que $gP_F x \in F$. Aussi, $x - P_F x \in F^\perp$ d'où $g(x - P_F x) \in F^\perp$, i.e., $gx - gP_F x \perp F$.

Appliquant le point précédent à g, P_F on obtient : $P_G P_F = P_F P_G$. On calcule : $(P_G P_F)^* = (P_G P_F)^2 = P_G P_F$ et $\operatorname{im}(P_G P_F) = G \cap F$ d'où $P_G P_F = P_{G \cap F}$. ☐

Corollaire. *Soient $f_1, f_2, \ldots \in \operatorname{End}(E)$ des endomorphismes auto-adjoints qui en plus* **commutent entre eux***. Alors ils sont simultanément diagonalisables dans une base orthonormée. Autrement dit, il existe une base orthonormée $\mathcal{B}$ dans laquelle toutes les matrices $M_{\mathcal{B}}(f_i)$ sont diagonale, i.e., les membres de $\mathcal{B}$ sont des vecteurs propres de chacun des f_i.*

3.5. **Diagonalisation d'une forme quadratique dans une base orthonormée.**

Soit q une forme quadratique et soit f un endomorphisme auto-adjoint tel que $q(x) = \langle x, f(x) \rangle$. On a vu que dans une base orthonormée q et f ont la même matrice. Donc dans une base orthonormée de vecteurs propres de f la matrice de q est diagonale et q est une combinaison linéaire de carrés.

Proposition (Réduction aux axes principaux)**.** *Pour toute forme quadratique q il existe une une base orthonormée dans laquelle la matrice de q est diagonale et q est une combinaison linéaire de carrés : $q(x) = \sum_1^n a_i x_i^2$. Les coefficients a_i sont les valeurs propres de l'endomorphisme auto-adjoint f associé ($q(x) = \langle x, f(x) \rangle$).*

On peut reformuler ce résultat comme la *diagonalisation simultanée de deux formes quadratiques :* la forme q et le produit scalaire $\langle x, x \rangle$.

3.6. **Endomorphismes positifs.** Si f est un endomorphisme symétrique, la forme bilinéaire $\varphi_f = \langle x, fy\rangle$ est symétrique (et coïncide avec $\langle fx, y\rangle$). On l'appelle la **forme bilinéaire associée** à f.

Un endomorphisme symétrique f est dit **positif** (respectivement, **défini positif**) si la forme associée $\langle x, fy\rangle$ est positive (respectivement, définie positive).

Le théorème spectral montre qu'un endomorphisme symétrique est positif (respectivement, défini positive) si et seulement si toutes ses valeurs propres sont positives (respectivement, stictement positives).

Une matrice symétrique A est dite **positive** (respectivement, **définie positive**) si ${}^tXAX \geq 0$ pour tout $X \in R^n$ (respectivement, ${}^tXAX > 0$ pour tout X non-nul).

.

Exemple : racine carré d'une matrice positive. Soit f un endomorphisme positif, Π_i le projecteur spectral associé à la valeur propre λ_i, $i = 1, ..., k$. On a $f = \sum \lambda_i \Pi_i$. Posons $g = \sum \sqrt{\lambda_i}\Pi_i$. Alors g est symétrique positif et $g^2 = f$. On montre facilement qu'une telle racine carré positive $\sqrt{f} = g$ est unique (en considérant la décomposition spectrale de g).

3.7. **Décomposition polaire. Théorème.** Soit f un endomorphisme inversible. Il existe un unique endomorphisme orthogonal U et un unique endomorphisme défini positif S tels que $f = US$.

Construction : On pose $S = \sqrt{f^* f}$; S est défini positif. On définit U par $U = fS^{-1}$ et on vérifie que U est orthogonal : $U^*U = S^{-1}f^*fS^{-1} = S^{-1}S^2S^{-1} = \mathrm{Id}$.

3.8. **Symétrie par rapport à un sous-espace. Réflexions.** Soit F un sous-espace de E et $E = F \oplus F^\perp$ la décomposition orthogonale. Pour $x \in E$ on écrit $x = x_1 + x_2$ avec $x_1 \in F$ et $x_2 \in F^\perp$. La **symétrie** s_F par rapport à F est définie par $s_F(x) = x_1 - x_2$.

Une symétrie par rapport à un hyperplan s'appelle **réflexion**. Si F est un hyperplan et e le vecteur unitaire orthogonal à F, la réflexion par rapport au F s'écrit $s_F(x) = x - 2\langle x, e\rangle e$.

Plus généralement, si z est un vecteur quelconque et $F = z^\perp$ alors $s_F(x) = x - 2\frac{\langle x,z\rangle}{\|z\|^2}z$.

Matrice d'une réflexion :

$$\begin{pmatrix} \mathrm{Id}_{n-1} & \\ & -1 \end{pmatrix}$$

Le déterminant d'une réflexion est donc toujours -1.

Lemme. *Toute réflexion est orthogonale.*

Théorème. *Soit $f \in \mathrm{End}(E)$ orthogonal, $r = \mathrm{rg}(f - \mathrm{Id})$. Alors f s'exprime comme produit d'au plus r réflexions. En particulier, f s'exprime comme produit d'au plus $n = \dim E$ réflexions.*

Démonstration. Par récurrence sur r.

Supposons que c'est vrai pour $r' < r$ et montrons pour r. Si $r = 0$: $f = \mathrm{Id}$ est la composition de zéro réflexions. Sinon, $f \neq \mathrm{Id}$, et il existe y tel que $fy \neq y$. Soit $z = fy - y \neq 0$ et $F = z^\perp$. Alors :

$\|z\|^2 = 2\|y\|^2 - 2\langle y, fy\rangle$.

Pour tout x : $s_F(x) = x - 2\frac{\langle x,z\rangle}{\|z\|^2}z = x - \frac{\langle x,z\rangle}{\|y\|^2 - \langle y,fy\rangle}z$.

D'où $s_F(y) = f(y)$ et $s_F(z) = -z$.

Si $fx = x$ alors $x \perp z$ et $s_F(x) = x$.

On applique l'hypothèse de récurrence à $s_F f$. □

Transformations orthogonales en petite dimension.

Dimension 1. $Ux = x$ ou $Ux = -x$.

Dimension 2. a) dét $U > 0$: U est une rotation.

Sa matrice est $U = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$.

b) dét $U < 0$: U est une réflexion.

Sa matrice est $U = \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}$.

.

Dimension 3. a) dét $U > 0$: U est une rotation autour d'un axe.

b) dét $U < 0$: U est une réflexion composée avec une rotation autour de l'axe orthogonal au plan de réflexion.

Remarquer qu'il y a (au moins) une valeur propre réelle.

3.9. **Réduction des endomorphismes orthogonaux.**

Lemme. *Soit $O \in M_n(\mathbf{R})$ une matrice orthogonale. Alors toutes les valeurs propres (complexes) de O sont de module* 1.

Lemme. *Soit $U \in M_n(\mathbf{C})$ tel que $U^{-1} = {}^t\overline{U}$ (une telle matrice est appelée* **unitaire***). Alors toutes les valeurs propres de U sont de module* 1.

Démonstration. Supposons que $X \in \mathbf{C}^n \setminus \{0\}$, $UX = \lambda X$, et posons $\|X\|^2 = {}^t\overline{X}X = \sum |x_i|^2 > 0$. Alors :

$$\|X\|^2 = {}^t\overline{X}X = {}^t\overline{X}({}^t\overline{U}U)X = {}^t\overline{(UX)}UX = (\bar{\lambda}{}^tX)(\lambda X) = \lambda\bar{\lambda}\|X\|^2,$$

d'où $|\lambda|^2 = \lambda\bar{\lambda} = 1$. □

Lemme. *Soit $A \in M_n(\mathbf{R})$ quelconque. Il existe un sous espace $F \subseteq \mathbf{R}^n$, $1 \leq \dim F \leq 2$, qui est stable par A.*

Démonstration. Si A admet une valeur propre réelle λ, il existe $X \in \mathbf{R}^n \setminus \{0\}$ tel que $AX = \lambda X$, donc $F = \mathrm{Vect}(X)$ est stable par A, de dimension 1.

Sinon, soit $\lambda \in \mathbf{C} \setminus \mathbf{R}$ une valeur propre, et soit $X \in \mathbf{C}^n$ un vecteur propre, $AX = \lambda X$, d'où $A\overline{X} = \bar{\lambda}\overline{X}$. Puisque $\lambda \neq \bar{\lambda}$, X et $\overline{X}$ ne sont pas colinéaire, on peu donc poser $Y = X + \overline{X} \in \mathbf{R}^n \setminus \{0\}$. An a $AY = \lambda X + \bar{\lambda}\overline{X}$ et $A^2Y = \lambda^2 X + \bar{\lambda}^2\overline{X} = (\lambda + \bar{\lambda})AY - \lambda\bar{\lambda}Y$, où $\lambda + \bar{\lambda}, \lambda\bar{\lambda} \in \mathbf{R}$. Donc $F = \mathrm{Vect}(Y, AY) \subseteq \mathbf{R}^n$ est stable par A. □

Proposition. *Soit U un endomorphisme orthogonal.*

(i) *Si le sous-espace F est stable par U, alors $F^\perp$ est stable par U.*

(ii) *Toute valeur propre de U est de module 1.*

(iii) *E se décompose en somme orthogonale des sous-espaces stables par U de dimension 1 ou 2.*

Démonstration. (i) On a $U(F) \subseteq F$ et puisque U est inversible, $\dim U(F) = \dim F$, d'où $F = U(F)$. Ainsi, F est stable par U^{-1}, i.e., par U^*. Maintenant, si $x \in F^\perp$ et $y \in F$ alors $U^*y \in F$ et $\langle Ux, y\rangle = \langle x, U^*y\rangle = 0$, donc $Ux \in F^\perp$ également.

(ii) D'après le Lemme.

(iii) D'après le Lemme précédent, E admet un sous-espace vectoriel $F \subseteq E$, tel que F est stable par U et $\dim F \leq 2$. Puisque $F^\perp$ est aussi stable par E on peut procéder par récurrence.

□

.

. .

4.7. Orthogonalisation de Gram-Schmidt et la décomposition orthogonale-triangulaire (décomposition QR).

Théorème. Soit A une matrice inversible. Il existe l'unique matrice orthogonale Q et l'unique matrice triangulaire supérieure R avec une diagonale positive telles que $A = QR$.

Construction : La matrice ${}^tAA$ est définie positive. Soit ${}^tAA = {}^tRR$ la décomposition de Cholesky, où R est triangulaire supérieure. On définit Q par $Q = AR^{-1}$ et on vérifie que Q est orthogonal : ${}^tQQ = {}^tR^{-1}.{}^tAAR^{-1} = {}^tR^{-1}.{}^tRRR^{-1} = Id$.

.

La décomposition QR est liée à l'orthogonalisation de Gram-Schmidt. Soit $v_1, ..., v_n$ les colonnes de A, $e_1, ..., e_n$ les colonnes de Q et $B = R^{-1}$ (B est triangulaire supérieure et $(e_1, ..., e_n$) est une base orthonormée).

La relation $A = QR$ est équivalente à $Q = AB$ qui s'écrit en termes des colonnes de façon suivante :

$e_1 = b_{11}v_1$,

$e_2 = b_{12}v_1 + b_{22}v_2$,

....

$e_k = \sum_{i=1}^k b_{ik}v_i$

Donc e_k est une combinaison linéaire des vecteurs $v_1, ..., v_k$ pour tout $k \geq 1$.

4. Espaces Hermitiens

1. Formes hermitiennes.

.

La théorie des formes et espaces hermitiens est parallèle à celle des formes bilinéaires symétriques et des epaces euclidiens.

1.1. Définition. Soit E un espace vectoriel sur C.

Une **forme hermitienne** sur E est une application $h : E \times E \to C$, vérifiant :

1. h est linéaire à droite : $h(x, \alpha y + \beta z) = \alpha h(x, y) + \beta h(x, z)$,

2. h est semi-linéaire à gauche : $h(\alpha x + \beta y, z) = \overline{\alpha} h(x, z) + \overline{\beta} h(y, z)$.

3. Symétrie hermitienne : $h(y, x) = \overline{h(x, y)}$ pour tous $x, y \in E$.

(Noter que 2. est la conséquence de 1. et 3.)

.

Pour une forme hermitienne h on définit la **forme quadratique associée** $q_h : E \to R : q_h(x) = h(x, x)$. Noter que q_h est à valeurs **réelles**.

La forme hermitienne est déterminée par la forme quadratique associée :

$$h(x, y) = \frac{1}{4}[q_h(x + y) - q_h(x - y) + iq_h(x - iy) - iq_h(x + iy)]$$

(*"identité de polarisation"*).

.

L'ensemble de toutes les formes hermitiennes est un espace vectoriel sur R : si $h_1, ..., h_k$ sont des formes hermitiennes et $a_1, ...a_k$ des scalaires réels, $a_1\varphi_1 + ... + a_k\varphi_k$ est une forme hermitienne.

.

Exemples. 1. Si f et g sont deux formes C-linéaires, $\varphi(x, y) = \overline{f(x)}g(y) + \overline{g(x)}f(y)$ est une forme hermitienne.

2. Soit E l'espace des matrices $k \times n$ sur C ; alors $h(A, B) = tr({}^t\bar{A}B)$ est une forme hermitienne.

3. Soit $E = C([a, b], C)$ l'espace des fonctions continues sur $[a, b]$. Alors $\varphi(f, g) = \int_a^b \overline{f(t)}g(t)dt$ est une forme hermitienne.

.

1.2. Expression en coordonnées. On suppose que dim $E = n < \infty$.

Soit $\mathcal{B} = (e_1, ..., e_n)$ une base de E, $x = \sum_1^n x_ie_i$, $y = \sum_1^n y_ie_i$.

Alors $h(x, y) = \sum_{i,j=1}^n \overline{x_i}y_j(e_i, e_j) = \sum_{i,j=1}^n a_{ij}\overline{x_i}y_j$ où $a_{ij} = h(e_i, e_j)$.

La matrice $A = (a_{ij}) = (h(e_i, e_j))$ est la *matrice de la forme hermitienne* h dans la base $\mathcal{B}$.

Cette matrice est **hermitienne** : $a_{ij} = \overline{a_{ji}}$, ou $\overline{{}^tA} = A$. La forme quadratique associée s'écrit : $q_h(x) = \sum_{i,j=1}^n a_{ij}\overline{x_i}x_j$.

Soit X la colonne des composantes du vecteur x : ${}^tX = (x_1, ..., x_n)$. Alors on peut écrire h à l'aide de la multiplication matricielle :

$$h(x,y) = \overline{{}^tX}AY$$

.

1.3. Changement de base (changement linéaire de coordonnées).

Soit $\mathcal{B}' = (e'_1, ..., e'_n)$ une autre base de E, soit X' et Y' les colonnes des coordonnées des vecteurs x et y dans la base $\mathcal{B}'$.

On a $X = PX'$ et $Y = PY'$, où P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Alors $h(x,y) = \overline{{}^tX}AY = \overline{{}^tX'{}^tP}APY' = \overline{{}^tX'}A'Y'$ où

$$A' = \overline{{}^tP}AP$$

est la matrice de la forme h dans la base $\mathcal{B}'$.

.

1.4. Equivalence des formes. Deux formes hermitiennes h et h' définies dans E et E' sont dites **équivalentes** si il existe un isomorphisme $f : E \to E'$ tel que $h(x,y) = h'(f(x), f(y))$.

Si $\dim(E) < \infty$, les formes h et h' sont équivalentes si leurs matrices A et B sont liées par $B = \overline{{}^tP}AP$ avec P inversible (autrement dit, si on peut trouver deux bases dans lequelles h et h' ont la même matrice).

.

1.5. On appelle **rang** d'une forme hermitienne le rang de sa matrice (il ne dépend pas du choix de la base). On dit que la forme est **non-dégénérée** si son rang est égal à la dimension de E.

Le **noyau** de h est défini par

Ker $h = \{x \in E : \forall y \in E, h(x,y) = 0\}$.

On a : rang (h) + dim (Ker h) = dim (E).

.

1.6. Soit h une forme hermitienne. Les vecteurs x et y sont **orthogonaux** si $h(x,y) = 0$. Une base est dite **orthogonale** si ses vecteurs sont deux à deux orthogonaux.

Dans une base orthogonale la forme s'écrit $h(x,y) = \sum_1^n a_i \overline{x_i} y_i$, avec $a_i \in R$ et la matrice de h est diagonale. La forme quadratique associée devient alors une combinaison linéaire de carrés : $q(x) = \sum_1^n a_i |x_i|^2$.

Le rang de h (ou de q) est le nombre de coefficients a_i non-nuls. Le noyau de h est engendré par les vecteurs de base e_i pour lesquels $a_i = 0$ (a_i réels).

.

1.7. Orthogonalisation de Gauss (réduction en carrés).

L'orthogonalisation de Gauss permet de fabriquer une base orthogonale pour la forme quadratique hermitienne $q_h(x) = \sum_{i,j=1}^n a_{ij} \overline{x_i} x_j$ par des changements de coordonnées successives.

La méthode est la même que pour les formes bilinéaires symétriques.

Résultat : toute forme hermitienne admet une base orthogonale.

.

1.8. Classification des formes hermitiennes. Signature

Soit q une forme quadratique hermitienne. Dans une base orthogonale, si on regroupe les coefficients positifs et négatifs, q s'écrit :

$q(x) = \sum_{i=1}^r a_i |x_i|^2 - \sum_{i=r+1}^{r+s} a_i |x_i|^2$ avec $a_i > 0$, $i = 1, ..., k$ et $r + s$ est le rang de q.

.

Théorème (loi d'inertie de Sylvester). Les entiers r et s (le nombre de carrés positif et négatifs) sont indépendants du choix de la base $q-$orthogonale.

.

Le couple (r, s) s'appelle **signature** de la forme hermitienne.

.

Corollaire. Deux formes hermitiennes sont équivalentes si et seulement si elles ont la même signature.

.

1.9. La forme quadratique hermitienne q est dite **positive** si $q(x) \geq 0$ pour tout $x \in E$ (donc, si $s = 0$) ; elle est dite **définie positive** si $q(x) > 0$ pour tout x non-nul (donc, si $r = dim(E)$).

.

En termes matriciels, A est positive si $\overline{{}^tX}AX \geq 0$ pour tout X ; A est définie positive si $\overline{{}^tX}AX > 0$ pour tout $X \neq 0$.

Remarque : pour toute matrice C la matrice $A = \overline{{}^tC}C$ est positive ; $\overline{{}^tC}C$ est définie positive si et seulement si C est inversible.

.

1.10. Orthogonalisation de Gauss pour les formes définie positives.

Si $q(x) = \sum_{i,j=1}^{n} a_{ij}\overline{x_i}x_j$ est définie positive, on a $a_{ii} > 0$ pour tout i. Donc dans l'algorithme de Gauss la matrice de changement de variables est à chaque étape triangulaire (supérieure) ; la matrice de passage P vers la base orthonormale dans laquelle q est la somme des carrés est donc triangulaire supérieure et $\overline{{}^tP}AP = I_n$. Soit $C = P^{-1}$. On a $A = \overline{{}^tC}C$.

Théorème de factorisation triangulaire (Gauss-Cholesky).

Pour toute matrice A hermitienne définie positive il existe une unique matrice C triangulaire supérieure à diagonale positive telle que $A = \overline{{}^tC}C$.

.

2. Espaces Hermitiens.

2.1. Soit E un C-espace vectoriel. Un **produit scalaire** sur E est une forme hermitienne définie positive, noté $<.,.>$.

La **norme** associée est définie par $\|x\|^2 = <x, x>$.

L'inégalité de Cauchy-Schwartz $|<x, y>| \leq \|x\|\|y\|$ entraine l'inégalité du triangle $\|x + y\| \leq \|x\| + \|y\|$.

La distance d dans E est définie par $d(x, y) = \|x - y\|$.

Le produit scalaire est déterminé par la norme associée ("identité de polarisation").

Un C-espace vectoriel de dimension finie muni d'un produit scalaire s'appelle **espace hermitien**.

Exemples : 1. Produit scalaire canonique dans C^n : $<x, y> = \sum_1^n \overline{x_i}y_i$; la norme est donnée par le "théorème de Pythagore" : $\|x\|^2 = \sum_1^n |x_i|^2$.

2. $E = C([a, b], C)$, $<f, g> = \int_a^b \overline{f(t)}g(t)dt$.

.

2.2. Deux vecteurs x et y sont **orthogonaux** si $<x, y> = 0$.

Sous-espace orthogonale. Soit $A \subset E$; **l'orthogonal** de A est l'ensemble de vecteurs de E orthogonaux à tous les vecteurs de A :

$A^{\perp} = \{x \in E : \forall y \in A \text{ on a } <x, y> = 0\}$.

Il est claire que $A^{\perp}$ est un sous-espace vectoriel de E.

Deux sous-espaces E_1 et E_2 sont **orthogonaux** si tout vecteur de E_1 est orthogonal à tout vecteur de E_2. Ceci est équivalent à dire que $E_2 \subset E_1^{\perp}$ ou que $E_1 \subset E_2^{\perp}$. Il est évident dans ce cas que $E_1 \bigcap E_2 = \{0\}$.

.

Famille orthogonale. Une famille de vecteurs de E est dite **orthogonale** si les vecteurs de cette famille sont deux à deux orthogonaux.

Une famille de vecteurs de E est dite **orthonormale** si elle est orthogonales et tous ses vecteurs sont de norme 1.

Lemme. Une famille orthogonale sans vecteurs nuls est libre.

D'une famille orthogonale $(e_1, ..., e_n, ...)$ on peut facilement passer à une famille orthonormale en normalisant les vecteurs e_i : $e'_i = \frac{e_i}{\|e_i\|}$.

.

Exemple. 1. Dans l'espace des fonctions continues $C([0, 2\pi], C)$ avec le produit scalaire $< f, g >= \frac{1}{2\pi}\int_0^{2\pi} \overline{f(t)}g(t)dt$ la famille $(e^{int})_{n\in Z}$ est orthonormale.

.

2.3. Coordonnées dans une base orthonormale.

Soit $(e_1, ..., e_n)$ une base **orthonormale** , soit $x = \sum_1^n x_i e_i$, $y = \sum_1^n y_i e_i$. Alors $< x, y >= \sum_1^n \overline{x_i} y_i$, $\|x\|^2 = \sum_1^n |x_i|^2$ ("théorème de Pythagore") et $x_i =< e_i, x >$.

Coordonnées dans une base **orthogonale :** $< x, y >= \sum_1^n < e_i, e_i > \overline{x_i} y_i$

$\|x\|^2 = \sum_1^n < e_i, e_i > |x_i|^2$ et $x_i = \frac{< e_i, x >}{< e_i, e_i >}$.

.

2.4. Orthogonalisation de Gram-Schmidt.

Soit $(v_1, ..., v_n, ...)$ une famille libre dans E. On peut construire une famille orthonormale $e_1, ..., e_n, ...$ telle que $Vect(v_1, ..., v_k) = Vect(e_1, ..., e_k)$ pour tout $k \geq 1$. (Autrement dit, e_k est une combinaison linéaire de $v_1, ..., v_k$.)

.

La construction est la même que pour un produit scalaire euclidien.

.

Corollaire. Tout espace hermitien admet une base orthonormale. Toute famille orthonormale dans un espace hermitien peut être complétée en une base orthonormale.

.

2.5. Projection orthogonale.

Soit E un espace muni du produit scalaire et $F \subset E$ un sous-espace.

Motivation. On sait que $F \bigcap F^\perp = \{0\}$. Supposons que $F^\perp$ est un supplémentaire de F, donc $E = F \bigoplus F^\perp$, somme directe orthogonale. (Ceci est toujours vrai en dimension finie.)

Le projecteur othogonal sur F, noté P_F, est par définition le projecteur sur F parallèlement à $F^\perp$. Si $x \in E$ on décompose $x = x_1 + x_2$ avec $x_1 \in F$ et $x_2 \in F^\perp$; alors, par définition, $P_F(x) = x_1$, la composante "orthogonale" de x dans F.

Noter que le projecteur orthogonal sur $F^\perp$ est $P_{F^\perp} = Id - P_F$.

.

Définition. Soit $F \subset E$ un sous-espace de dimension finie.

Soit $(e_1, ..., e_n)$ une base **orthonormale** de F.

On définit $P_F : E \to E$ par $P_F(x) = \sum_1^n < e_i, x > e_i$. Alors on a

Lemme. P_F est un projecteur sur F parallèlement à $F^\perp$.

Corollaire. Si F est un sous-espace de dimension finie, $F^\perp$ est un supplémentaire de F : $E = F \bigoplus F^\perp$, somme directe orthogonale. On a aussi $(F^\perp)^\perp = F$.

.

Projection orthogonale dans une base quelconque.

Le vecteur $y = P_F(x)$ est caractérisé par les conditions

$y \in F$ et $< z, y >=< z, x >$ pour tout vecteur z de F.

Soit $(e_1, ..., e_n)$ une base de F. La dernière condition est donc équivalente à $< e_j, y >=< e_j, x >$, $j = 1, ..., n$.

Posons $P_F(x) = \sum_1^n y_i e_i$. pour déterminer les coefficients y_i on doit résoudre le système :

$\sum_{j=1}^n y_j < e_i, e_j >=< e_i, x >$, $j = 1, ..., n$.

La matrice de ce système $G = (< e_i, e_j >)$ s'appelle *matrice de Gram.* Soit $(\tilde{e}_1, ..., \tilde{e}_n)$ une base orthonormale et $e_j = \sum_i a_{ij}\tilde{e}_i$, donc

$a_{ij} = (< e_i, \tilde{e}_j >)$. Soit $A = (a_{ij})$, alors $G = \overline{{}^t A} A$. En particulier, G est définie positive et dét G=|dét $A|^2$.

.

2.6. Projection othogonale et meilleure approximation en moyenne quadratique. Distance à un sous-espace.

Lemme. Soit F est un sous-espace de dimension finie et $x \in E$. Alors la projection $P_F(x)$ réalise la distance minimale entre x et les vecteurs de F : $\|x - P_F(x)\| = \min \{\|x - z\|, z \in F\}$.

.

Exemple. Meilleure approximation en moyenne quadratique par des polynômes trigonométriques.

Un polynôme trigonométrique de dégré $\leq n$ est la somme $p(t) = \sum_{k=-n}^{n} a_k e^{ikt}$. Soit $f \in C([0, 2\pi])$ une fonction continue. On cherche on polynôme trigonométrique p de degré $\leq n$ tel que l'écart $\int_0^{2\pi} |f(t) - p(t)|^2 dt$ soit minimal.

La réponse est donnée par la projection orthogonal dans $C([0, 2\pi])$ sur le sous-espace des polynômes trigonométriques de dégré $\leq n$; le produit scalaire est $< f, g >= \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)} g(t) dt$. On a la famille orthogonale : $(e^{int})_{n \in Z}$. On en déduit les coefficients du polynôme $p(t)$ de meilleure approximation : $a_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikt} f(t) dt$.

.

2.7. Inégalité de Bessel et égalité de Bessel-Parseval.

Soit E un espace hermitien, soit $(e_1, ..., e_n)$ une famille orthonormale.

Soit $x \in E$ et $x_i =< e_i, x >$. Il est claire que $\sum_{i=1}^{n} |x_i|^2 \leq \|x\|^2$.

.

Lemme. Soit $x \in E$. Les propriétés suivantes sont équivalentes :

(i) $\sum_{i=1}^{n} |x_i|^2 = \|x\|^2$.

(ii) $x = \sum_{i=1}^{n} x_i e_i$.

(iii) x appartient à l'espace vectoriel engendré par $(e_1, ..., e_n)$ (donc x est une cobinaison linéaire des vecteurs $e_1, ..., e_n$).

.

Soit maintenant $\dim(E) = \infty$.

Théorème. Soit $(e_1, ..., e_n, ...)$ une famille orthonormale infinie, $x \in E$ et $x_i =< e_i, x >$.

A. Pour tout n on a $\sum_{i=1}^{n} |x_i|^2 \leq \|x\|^2$

et la série $\sum_{i=1}^{\infty} |x_i|^2$ converge.

B. Les propriétés suivantes sont équivalentes :

(i) $\sum_{i=1}^{\infty} |x_i|^2 = \|x\|^2$.

(ii) $x = \lim_{n \to \infty} \sum_{i=1}^{n} x_i e_i$.

(iii) x appartient à l'adhérence de l'espace vectoriel engendré par la suite $(e_1, ..., e_n, ...)$.

.

Egalité de Bessel-Parseval dans une "base" orthogonale.

Si $(e_1, ..., e_n, ...)$ est une famille orthogonale et x appartient à l'adhérence de l'espace vectoriel engendré par la suite $(e_1, ..., e_n, ...)$, alors

$\|x\|^2 = \sum_{i=1}^{\infty} \frac{<e_i, x>^2}{<e_i, e_i>}$.

.

3. Formes hermitiennes et endomorphismes.

3.1. Proposition. Soit E un espace hermitien et f un endomorphisme de E. Il existe l'unique endomorphisme f^* tel que

$< x, f(y) >=< f^*(x), y >$.

Définition. L'endomorphisme f^* tel que $< x, f(y) >=< f^*(x), y >$ s'appelle **l'adjoint** de f.

La matrice de f^* dans une base orthonormale est la transposée hermitienne de la matrice de f : $M_{f^*} = \overline{{}^t M_f}$.

Définition. L'endomorphisme f est dit **auto-adjoint** ou **hermitien** si $f^* = f$. L'endomorphisme f est hermitien si et seulement si sa matrice dans une base orthonormale est hermitienne.

.

Exemple. Une projection orthogonale est auto-adjoint.

.

3.2. *Propriétés de l'adjoint.* L'application $f \to f^*$ est linéaire ; $(f^*)^* = f$, $(fg)^* = g^*f^*$ et $(f^{-1})^* = (f^*)^{-1}$.

Lemme. (1) L'orthogonal d'un sous-espace stable par f est stable par f^*.

(2) Ker $f^* =$ (Im $f)^\perp$ et Im $f^* =$(Ker $f)^\perp$.

.

Corollaire. Si f est auto-adjoint, alors Ker $f =$ (Im $f)^\perp$ et E est la somme orthogonale de Ker f et Im f. L'orthogonal d'un sous-espace stable par f est stable par f.

.

3.3. Diagonalisation des matrices hermitiennes.

Proposition. Soit f un endomorphisme auto-adjoint. Alors

(i) Toutes les valeurs propres de f sont réelles.

(ii) Les sous-espaces propres de f sont deux à deux orthogonaux.

(iii) f est diagonalisable dans une base orthonormale.

.

3.4. Un endomorphisme hermitien f est dit **positif** (respectivement, **défini positif**) si la forme associée $h(x, y) =< x, f(y) >$ est positive (respectivement, défini positive).

La proposition précédente montre que'un endomorphisme hermitien est positif (respectivement, défini positive) si et seulement si toutes ses valeurs propres sont positives (respectivement, stictement positives).

Une matrice hermitienne A est dit **positive** (respectivement, **défini positive**) si $\overline{^tX}AX \geq 0$ pour tout $X \in R^n$ (respectivement, $\overline{^tX}AX > 0$ pour tout X non-nul).

.

Exemple : racine carré d'une matrice positive. Soit f un endomorphisme positif, Π_i le projecteur spectral associé à la valeur propre λ_i, $i = 1, ..., k$. On a $f = \sum \lambda_i \Pi_i$. Posons $g = \sum \sqrt{\lambda_i}\Pi_i$. Alors g est hermitien positif et $g^2 = f$. On montre facilement qu'une telle racine carré positive $\sqrt{f} = g$ est unique.

.

3.5. Diagonalisation d'une forme hermitienne dans une base orthonormale.

Soit q une forme quadratique hermitienne et soit f un endomorphisme auto-adjoint tel que $q(x) =< x, f(x) >$. On a vu que dans une base orthonormale q et f ont la même matrice. Donc dans une base orthonormale de vecteurs propres de f la matrice de q est diagonale et q est une combinaison linéaire de carrés.

Proposition. *"Réduction aux axes principaux".* Pour toute forme quadratique hermitienne q il existe une une base orthonormale dans laquelle la matrice de q est diagonale et q est une combinaison linéaire de carrés : $q(x) = \sum_1^n a_i|x_i|^2$. Les coefficients a_i sont les valeurs propres de l'endomorphisme auto-adjoint f associé ($q(x) =< x, f(x) >$).

On peut reformuler ce résultat comme la *diagonalisation simultanée de deux formes quadratiques :* la forme q et le produit scalaire $< x, x >$.

.

4. Transformations unitaires.

4.1. Soit E un espace hermitien.

Un endomorphisme U de E est **unitaire** (est une isométrie linéaire) si U préserve le produit scalaire : $< Ux, Uy >=< x, y >$ pour tout $x, y \in E$.

.

Proposition. Les propriétés suivantes sont équivalentes.
(i) U est unitaire.
(ii) U préserve la norme : $\|Ux\| = \|x\|$ pour tout $x \in E$.
(iii) U transforme toute base orthonormale en base orthonormale.
(iv) U transforme une base orthonormale en base orthonormale.
Un endomorphisme unitaire est injectif, donc inversible
(dim $E < \infty$!). Son inverse est aussi unitaire.

.

4.2. L'égalité $< Ux, Uy >=< x, y >$ s'écrit aussi
$< U^*Ux, y >=< x, y >$ ce qui est équivalent à $U^*U = I_n$. Donc U est orthogonal si et seulement si $U^*U = I_n$, ou encore ssi $U^{-1} = U^*$.

Une **matrice unitaire** est la matrice d'un endomorphisme unitaire dans une base orthonormale. Une matrice unitaire U est caractérisée par la relation $\overline{{}^tU}U = U\overline{{}^tU} = I_n$ qui signifie que les colonnes de A (aussi que ses lignes) constituent une base orthonormale par rapport au produit scalaire canonique dans C^n.

.

4.3. *Transformations unitaires et diagonalisation d'une forme hermitienne dans une base orthonormale.*

Soit A la matrice d'une forme quadratique hermitienne q dans une base $\mathcal{B}$; soit P la matrice de passage de $\mathcal{B}$ à une autre base $\mathcal{B}'$. Alors la matrice de la forme q dans la base $\mathcal{B}'$ est $A' = \overline{{}^tP}AP$. Si les deux bases $\mathcal{B}$ et $\mathcal{B}'$ sont orthonormales, P est unitaire : $\overline{{}^tP} = P^{-1}$ et on a $A' = \overline{{}^tP}AP = P^{-1}AP$. Donc la matrice d'une forme se transforme comme la matrice d'un endomorphisme si le changement de coordonnées est unitaire. Cela montre encore une fois que la diagonalisation d'une forme quadratique par une transformation unitaire demande la recherche des valeurs et des vecteurs propres de A.

.

.

4.4. Réduction des endomorphismes unitaires.
Proposition. Soit U un endomorphisme unitaire.
(i) Si le sous-espace F est stable par U, alors $F^\perp$ est stable par U.
(ii) Toute valeur propre de U est de module 1.
(iii) U est diagonalisable dans une base orthonormale.
Transformations unitaires en petite dimension.
Dimension 1. $Ux = \alpha x$ où $|\alpha| = 1$.
Dimension 2.
Toute matrice unitaire peut s'écrire sous la forme suivante :
$U = \alpha \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}$ où $|\alpha| = 1$ et $|a|^2 + |b|^2 = 1$.

.

4.5. Décompositon polaire.
Théorème. Soit $A \in Mat_n(C)$ une matrice inversible. Il existe l'unique matrice unitaire U et l'unique matrice hermitienne définie positive S telles que $A = US$.

.

4.6. Décomposition unitaire-triangulaire.
Théorème. Soit $A \in Mat_n(C)$ une matrice inversible. Il existe l'unique matrice unitaire U et l'unique matrice triangulaire supérieure T avec une diagonale positive telles que $A = UT$.

.

La décomposition unitaire-triangulaire est liée à l'orthogonalisation de Gram-Schmidt.

5. Géométrie affine

5.1. **Généralités.**

Définition. Soit E un $\mathbf{R}$-espace vectoriel de dimension finie.

Un **espace affine** dirigé par E est un ensemble non-vide $\mathcal{E}$ muni d'une application $\mathcal{E}\times\mathcal{E}\to E$ qui associe au couple de points $A,B\in\mathcal{E}$ un vecteur $\overrightarrow{AB}$, et qui vérifie

(i) Pour tout $O\in\mathcal{E}$ l'application $A\mapsto\overrightarrow{OA}$ est une bijection de $\mathcal{E}$ sur E.

(ii) Pour tous $A,B,C\in\mathcal{E}$ on a la *relation de Chasles :* $\overrightarrow{AC}=\overrightarrow{AB}+\overrightarrow{BC}$.

La *dimension* de $\mathcal{E}$ est par définition celle de E.

La relation de Chasles a pour conséquence immédiate : $\overrightarrow{AA}=\overrightarrow{0}$ et $\overrightarrow{BA}=-\overrightarrow{AB}$.

Example. L'espace vectoriel E admet une **structure canonique d'espace affine dirigé par** E : pour deux vecteurs u et v on pose $\overrightarrow{uv}=v-u$.

Fixons $O\in\mathcal{E}$ que l'on appellera « origine, » et posons $v(A)=\overrightarrow{OA}$. Alors $v:\mathcal{E}\to E$ est une bijection qui identifie la structure affine de $\mathcal{E}$ avec la structure affine canonique de E : $\overrightarrow{v(A)v(B)}=v(B)-v(A)=\overrightarrow{OB}-\overrightarrow{OA}=\overrightarrow{AB}$ - conséquence de la relation de Chasles. Le choix d'origine donc "vectorialise" l'espace affine. Réciproquement, on peut dire qu'un espace affine est un espace vectoriel avec l'origine (zéro) « oubliée ».

Définition (Translation)**.** Soit $v\in E$. Pour $A\in\mathcal{E}$ on définit $T_v(A)\in\mathcal{E}$ comme étant l'unique point tel que $\overrightarrow{AT_v(A)}=v$. On appelle T_v la **translation** par v.

Lorsque $\mathcal{E}=E$ (avec la structure affine canonique) on a $T_v(u)=v+u$. Par analogie, on adopte la notation : $T_v(a)=A+v$. Dans cette notation $B=A+\overrightarrow{AB}$ (et $v=\overrightarrow{AB}$ est l'unique tel que $B=A+v$).

Lemme. *On a :*

(i) $T_0=\mathrm{Id}$ *et* $T_{v+u}=T_vT_u$. *Il en suit que* $T_v\colon\mathcal{E}\to\mathcal{E}$ *est une bijection, avec* $T_v^{-1}=T_{-v}$. *(On dit alors que le groupe additif* $(E,+)$ **agit** *sur l'espace affine* $\mathcal{E}$*.)*

(ii) *Pour* $A,B\in\mathcal{E}$*, il existe un unique vecteur* $v\in E$ *tel que* $T_v(A)=B$ *(c'est* $v=\overrightarrow{AB}$*).*

Remarque. La donnée de la structure affine sur $\mathcal{E}$ est équivalente à la donnée de l'action de E sur $\mathcal{E}$ - la donnée d'une famille de "translations" T_v, $v\in E$, vérifiant les propriétés du lemme.

5.2. **Barycentres, combinaisons affines.**

Commençons avec un exemple :

Soit $v\in E$ et $A\in\mathcal{E}$. La **droite affine** passant par A dirigée par v est l'ensemble de points $M=A+\lambda v$, $\lambda\in\mathbf{R}$. La droite affine passant par deux points A et B est dirigée par le vecteur $v=\overrightarrow{AB}$; ses points sont $M=A+\lambda\overrightarrow{AB}$, $\lambda\in\mathbf{R}$.

Choisissons une "origine" O. Alors $M=A+\lambda\overrightarrow{AB}$ si et seulement si $\overrightarrow{OM}=\overrightarrow{OA}+\lambda\overrightarrow{AB}=(1-\lambda)\overrightarrow{OA}+\lambda\overrightarrow{OB}$. Autrement dit, la droite affine passant par A et B consiste en tous les points M tels que $\overrightarrow{OM}=\mu\overrightarrow{OA}+\lambda\overrightarrow{OB}$ avec $\lambda+\mu=1$. On appelle M le **barycentre** des points A et B affectés de poids μ et λ.

Lemme. *Soit* $A_1,\dots,A_k$ *des points affectés des poids* $\lambda_1,\dots,\lambda_k$*, tels que* $\sum_i\lambda_i\neq 0$. *Soit aussi* $O\in\mathcal{E}$ *un point quelconque (une « origine »). Les propriétés suivantes sont équivalentes pour un point* $M\in\mathcal{E}$ *:*

(i) $\sum_i\lambda_i\overrightarrow{MA_i}=0$.

(ii) $(\sum_i\lambda_i)\overrightarrow{OM}=\sum_i\lambda_i\overrightarrow{OA_i}$.

(iii) $\overrightarrow{OM} = \frac{1}{\sum_i \lambda_i} \sum_i \lambda_i \overrightarrow{OA_i}$.

(iv) $M = O + \frac{1}{\sum_i \lambda_i} \sum_i \lambda_i \overrightarrow{OA_i}$. *En particulier, un unique M ayant ces propriétés existe.*

De surcroît, un unique tel point M existe toujours.

Démonstration. (i) $\iff$ (ii). Puisque

$$\sum_i \lambda_i \overrightarrow{OA_i} - \left(\sum_i \lambda_i\right) \overrightarrow{OM} = \sum_i \lambda_i \left(\overrightarrow{OA_i} - \overrightarrow{OM}\right) = \sum_i \lambda_i \overrightarrow{MA_i} = 0.$$

(ii) $\iff$ (iii) $\iff$ (iv). Clair. $\square$

Définition. Le point M défini dans le lemme s'appelle le **barycentre** des points $A_1, ..., A_k$ affectés des coefficients $\lambda_1, ..., \lambda_k$. On notera $M = Bar\{(\lambda_i, A_i)\}$.

Si $M = Bar\{(\lambda_i, A_i)\}$ et $\sum_i \lambda_i = 1$, on écrit aussi $M = \sum_i \lambda_i A_i$.

En d'autres mots, $M = \sum_i \lambda_i A_i$ si et seulement si $\sum \lambda_i = 1$ et $\overrightarrow{OM} = \sum_i \lambda_i \overrightarrow{OA_i}$ (et ceci quel que soit O).

Si $\mathcal{E} = E$, le barycentre v des vecteurs $u_1, ..., u_k$ affectés des poids $\lambda_1, ..., \lambda_k$ est $v = \frac{1}{\sum_i \lambda_i} \sum_i \lambda_i u_i$. En particulier, si $\sum_i \lambda_i = 1$, alors $v = \sum_i \lambda_i u_i$. (Comparer avec la notation pour le barycentre : $M = \sum_i \lambda_i A_i$.)

Vu que v est une combinaison linéaire des vecteurs u_i, on peut appeler le barycentre « combinaison affine des points A_i ».

Définition (Repère affine)**.** On appelle **repère affine** de $\mathcal{E}$ la donnée de $n+1$ points $(A_0, \dots, A_n)$ tels que chaque $B \in \mathcal{E}$ s'exprime d'une manière unique comme $B = \sum_0^n x_i A_i$ (où $\sum_0^n x_i = 1$!). On appelle $(x_0, \dots, x_n)$ les **coordonnées barycentriques** de B dans ce repère.

Cette définition est symétrique, dans le sens qu'une permutation d'un repère est un repère. Brisons cette symétrie en donnant à A_0 un rôle particulier et posant $e_i = \overrightarrow{A_0A_i}$. On a donc $B = \sum_0^n x_i A_i$ si et seulement si $\overrightarrow{A_0B} = \sum_0^n x_i \overrightarrow{A_0A_i} = \sum_1^n x_i \overrightarrow{A_0A_i}$. Ainsi, $(A_0, \dots, A_n)$ est un repère si et seulement si chaque vecteur $v \in E$ s'exprime d'une manière unique comme $\sum_1^n x_i e_i$. Autrement dit :

Lemme. *Une famille de $n+1$ vecteurs $(A_0, \dots, A_n)$ est un repère affine si et seulement si $(e_1, \dots, e_n)$ est une base de E. En particulier, $n = \dim E$.*

Définition. Les **coordonnées affines** d'un point B dans le repère $(A_0, \dots, A_n)$ sont par définition les coordonnées du vecteur $\overrightarrow{A_0B}$ dans la base $(e_1, \dots, e_n)$.

Le lien : les coordonnées barycentriques de B dans un repère $(A_0, \dots, A_n)$ sont $(x_0, x_1, \dots, x_n)$ si et seulement si $(x_1, \dots, x_n)$ sont les coordonnées de $\overrightarrow{A_0B}$ dans la base $(e_1, \dots, e_n)$ et $x_0 = 1 - \sum_1^n x_i$.

La propriété suivante de "l'associativité du barycentre" est très utile.

Proposition. *Pour chaque $i = 1, \dots, k$ soit $B_i = \sum_{j=1}^m \lambda_{i,j} A_{i,j}$, où $\sum_j \lambda_{i,j} = 1$. Soit $C = \sum_{i=1}^k \mu_i B_i$, où $\sum \mu_i = 1$. Alors $\sum_{i,j} \mu_i \lambda_{i,j} = 1$, et $C = \sum_{i,j} \mu_i \lambda_{i,j} A_{i,j}$.*

Démonstration. On a

$$\sum_j \lambda_{i,j} \overrightarrow{B_iA_{i,j}} = 0, \qquad \sum_i \mu_i \overrightarrow{CB_i} = 0.$$

Ainsi

$$0 = \sum_i \mu_i \left(\sum_i \lambda_{i,j}\right) \overrightarrow{CB_i} + \sum_i \mu_i \sum_j \lambda_{i,j} \overrightarrow{B_iA_{i,j}} = \sum_{i,j} \mu_i \lambda_{i,j} (\overrightarrow{CB_i} + \overrightarrow{B_iA_{i,j}}) = \sum_{i,j} \mu_i \lambda_{i,j} \overrightarrow{CA_{i,j}}.$$

$\square$

5.3. **Sous-espace affine.**

Définition. Une partie non vide $\mathcal{F}$ de l'espace affine $\mathcal{E}$ est un **sous-espace affine** si pour tout $A, B \in \mathcal{F}$, $\mathcal{F}$ contient la droite passant par ces deux points.

Lemme. *Les conditions suivantes sont équivalentes pour une partie $\mathcal{F} \subseteq \mathcal{E}$:*

(i) *$\mathcal{F}$ est un sous-espace affine de $\mathcal{E}$.*

(ii) *Si $A_1, ..., A_k \in \mathcal{F}$ alors tout barycentre des points $A_1, ..., A_k$ est dans $\mathcal{F}$. (Toute combinaison affine des points de $\mathcal{F}$ est dans $\mathcal{F}$.)*

(iii) *Pour tout $O \in \mathcal{F}$ l'ensemble $F = \{\overrightarrow{OM}, M \in \mathcal{F}\}$ est un sous-espace vectoriel de E. De cette façon, $\mathcal{F}$ est à son tour un espace affine dirigé par F.*

(iv) *On a $\mathcal{F} = \{O + v, v \in F\}$ où F est un sous-espace vectoriel de E et O un point de $\mathcal{F}$.*

Une *droite* est un sous-espace affine de dimension 1.
Un *hyperplan* est un sous-espace affine de codimension 1 (de dimension $\dim \mathcal{E} - 1$).

Lemme. *L'intersection d'une famille de sous-espaces affines de $\mathcal{E}$ et un sous-espace affine* **si cette intersection est non vide**.

Soit $\mathcal{F}$ un sous-espace affine dirigé par F ; soit $O \in \mathcal{F}$ et $(v_1, ..., v_k)$ une base de F. On a donc un repère affine de $\mathcal{F}$ et tout point B de $\mathcal{F}$ admet l'expression unique : $B = O + \sum_{i=1}^{k} s_i v_i$, $(s_1, ..., s_k) \in R^k$.

Soit $(x_1, ..., x_n)$ un système de coordonnées affines dans $\mathcal{E}$. Soit $(c_1, ..., c_n)$ les coordonnées du point O et $(a_{1i} ..., a_{ni})$ les coordonnées du vecteur v_i. Alors les coordonnées du point $B = O + \sum_{i=1}^{k} s_i v_i$ sont $x_j = c_j + \sum_{i=1}^{k} a_{ji} s_i$, $j = 1, ..., n$.

Position relative de deux sous-espaces

Lemme. (i) *Soit $\mathcal{F}$ et $\mathcal{G}$ deux sous-espaces de $\mathcal{E}$ d'intersection non-vide. Alors $\mathcal{F} \cap \mathcal{G}$ est dirigé par $F \cap G$.*

(ii) *Si $F + G = E$ alors $\mathcal{F} \cap \mathcal{G}$ est non vide.*

(iii) *Si les sous-espaces vectoriels F et G sont supplémentaires ($E = F \oplus G$), alors $\mathcal{F} \cap \mathcal{G}$ est un seul point.*

Définition. On dit que deux sous-espaces $\mathcal{F}$ et $\mathcal{G}$ de $\mathcal{E}$ sont **parallèles** si $F \subset G$ ou $G \subset F$. Dans ce cas que soit $\mathcal{F}$ et $\mathcal{G}$ sont disjoints, soit l'un contient l'autre.

Sous-espace affine engendré par une partie.

Soit $S \subset \mathcal{E}$ une partie non-vide. Le sous-espace affine engendré par S, noté Aff(S), est l'intersection de tous les sous-espaces affines qui contiennent S.

Lemme. (i) *Aff(S) est l'ensemble de toutes les combinaisons affines (barycentres) des points de S.*

(ii) *Soit $O \in S$ et soit F le sous-espace vectoriel de $\mathcal{E}$ engendré par les vecteurs $\overrightarrow{OA}$, $A \in S$. Alors l'espace directeur de Aff(S) est F : Aff(S) $= \{O + v, v \in F\}$.*

(iii) *$(A_0, ..., A_n)$ est un repère affine si et seulement si $n = \dim \mathcal{E}$ et Aff$(A_0, ..., A_n) = \mathcal{E}$.*

5.4. **Applications affines.**

Définition. Une application $f : \mathcal{E} \to \mathcal{F}$ est **affine** si $f(\lambda A + \mu B) = \lambda f(A) + \mu f(B)$ pour tout $A, B \in \mathcal{E}$ et tout λ, μ tels que $\lambda + \mu = 1$.

Lemme. *Les conditions suivantes sont équivalentes pour une application $f : \mathcal{E} \to \mathcal{F}$:*

(i) *f est une application affine.*

(ii) *f préserve les barycentres : $f(Bar\{(\lambda_i, A_i)\}) = Bar\{(\lambda_i, f(A_i))\}$.*

(iii) *Pour un point $A \in \mathcal{E}$ l'application $h : E \to F$ défini par $h(\overrightarrow{AB}) = \overrightarrow{f(A)f(B)}$ ou, avec une autre notation, $h(v) = \overrightarrow{f(A)f(A+v)}$ est linéaire. [Ecriture équivalente : $f(A+v) = f(A) + h(v)$.]*

(iv) *Pour tout point A on a $f(A+v) = f(A) + h(v)$ où $h : E \to F$ est une application linéaire, la même pour tout A.*

(v) *f est différentiable et sa différentielle $df : E \to F$ est constante sur $\mathcal{E}$ (et alors $f(A+v) = f(A) + df(v)$.*

On peut dire qu'une application affine est une application linéaire plus une constante.

Ecriture en coordonnées. Fixons les repères affines $R = (A_0; e_1, ..., e_n)$ dans $\mathcal{E}$ et $R' = (B_0; e'_1, ..., e'_k)$ dans $\mathcal{F}$.

Soit $(x_1, ..., x_n)$ les coordonnées du point M dans R et $(y_1, ..., y_k)$ les coordonnées de $f(M)$ dans R'.

Alors $y_i = c_i + \sum_{j=1}^{n} a_{ij} x_j$, où (a_{ij}) est la matrice de la différentielle df dans les bases des repères R, R' et $(c_1, ..., c_k)$ les coordonnées du point $f(A_0)$.

Fonction affine $f : \mathcal{E} \to R$. Dans un système de coordonnées affines f s'écrit $f(x_1, ..., x_n) = c + \sum_i a_i x_i$.

.

La *formule du rang* pour les applications linéaires s'adapte aux applications affines : Soit $B \in Im(f)$. Alors dim $(\mathcal{E})$ =dim (Imf)+ dim $(f^{-1}(B))$.

Lemme. *Soit $f : \mathcal{E} \to \mathcal{F}$ une application affine. Alors*

(i) *Si $\mathcal{E}_1$ un sous-espace affine de $\mathcal{E}$ alors $f(\mathcal{E}_1)$ est un sous-espace affine de $\mathcal{F}$.*

(ii) *Si $\mathcal{F}_1$ un sous-espace affine de $\mathcal{F}$ et $f^{-1}(\mathcal{F}_1)$ est non-vide, alors, $f^{-1}(\mathcal{F}_1)$ est un sous-espace affine de $\mathcal{E}$.*

Equation d'un hyperplan : tout hyperplan est défini par une équation affine $\sum_i a_i x_i = c$.

.

"Combinaison affine" des applications affines.

Soit $f_i : \mathcal{E} \to \mathcal{F}$, $i = 1, ..., k$ des applications affines et $\lambda_1 + ... + \lambda_k = 1$.

On définit $f = \sum_i \lambda_i f_i$ par $f(A) = \sum_i \lambda_i f_i(A)$ (barycentre).

Exercice : définir une structure de l'espace affine dans l'ensemble de toutes les applications affines $f : \mathcal{E} \to \mathcal{F}$.

Applications affines de $\mathcal{E}$ dans $\mathcal{E}$. Transformations affines.

.

Les propriérés d'une application affine f sont déterminées par les propriétés de sa différentielle df.

Proposition. (i) *Si df n'admet pas de vecteur fixe non-nul (i.e., si 1 n'est pas une valeur propre de df), alors f admet un unique point fixe.*

(ii) *Si df admet des vecteurs fixes non-nuls (i.e., si 1 est une valeur propre de df), alors l'ensemble des points fixes de f est soit vide soit un sous-espace affine dirigé par $\{v \in E : df(v) = v\}$ (l'espace propre correspondant à 1).*

Si f admet un point fixe O et on prend $O = f(O)$ pour origine, f s'écrit $f(O+v) = O + df(v)$: donc, en vectorialisant $\mathcal{E}$ en O, on identifie f avec l'application linéaire df.

Définition. f est une **homothétie** de centre O et de rapport k si $f(O) = O$ et $df = kId$.

Proposition. (i) *$f : \mathcal{E} \to \mathcal{E}$ est une translation si et seulement si $df = \mathrm{Id}$.*

(ii) *Si $df = kId$ avec $k \neq 1$, alors f est une homothétie.*

Projections et symétries affines. Soit $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$, F l'espace vectoriel associé, et soit G un sous-espace vectoriel de E supplémentaire à F : $E = F \oplus G$.

.

Soit $\Pi : E \to E$ le projecteur vectoriel sur F parallèlement à G : $\Pi(v) = v$ si $v \in F$ et $\Pi(v) = 0$ si $v \in G$.

Définition. Soit $O \in F$. La **projection affine** p sur $\mathcal{F}$ parallèlement à G est définie par $p(A) = O + \Pi(\overrightarrow{OA})$.

Lemme. (i) *La projection affine sur $\mathcal{F}$ parallèlement à G ne dépend par du choix du point O. Ainsi, elle vérifie $p(A) = A$ pour tout $A \in \mathcal{F}$ et $dp = \Pi$, et son image est $\mathcal{F}$.*

(ii) *Soit f une transformation affine de $\mathcal{E}$. Alors f est un projecteur affine si et seulement si $f^2 = f$. Dans ce case c'est le projecteur affine sur $\mathcal{F} = f(\mathcal{E})$, parallèlement à $\ker df$.*

Soit $\sigma : E \to E$ la symétrie vectorielle par rapport à F parallèlement à G : $\sigma(v) = v$ si $v \in F$ et $\sigma(v) = -v$ si $v \in G$.

2.7. Définition. La **symétrie affine** s par rapport à $\mathcal{F}$ parallèlement à G est définie par les conditions $s(O) = O$ si $O \in \mathcal{F}$ et $ds = \sigma$; donc $p(M) = O + \sigma(\overrightarrow{OM})$.

Lemme. s est une symétrie affine si et seulement si $f \cdot f = Id$.

3. Espaces affines euclidiens

3.1. Définition. Un espace affine $\mathcal{E}$ est dit **euclidien** s'il est dirigé par une espace vectoriel euclidien E.

On définit la distance dans $\mathcal{E}$ par $d(A, B) = \|\overrightarrow{AB}\|$.

3.2. Définition. Un repère affine $(A_0, a_1, ..., A_n)$ est dit **orthonormé** si la base vectorielle $(\overrightarrow{A_0A_1}, ..., \overrightarrow{A_0A_n})$ est orthonormée.

3.3. Définition. Deux sous-espaces affine $\mathcal{F}$ et $\mathcal{G}$ sont **orthogonaux** si leurs espaces directeurs F et G sont orthogonaux.

.

Isométries

3.4. Définition. Soit X un espace métrique. On appelle **isométrie** de X une bijection $f : X \to X$ qui préserve la distance : $d(f(x), f(y)) = d(x, y)$ pour tous $x, y \in X$.

Remarque : les isométris de X forment un groupe par rapport à la composition.

Lemme. *Une application affine $f : \mathcal{E} \to \mathcal{E}$ est une isométrie si et seulement si sa partie linéaire $df : E \to E$ est une isométrie vectorielle.*

Lemme. *Une isométrie $f : E \to E$ qui fixe 0 est linéaire. C'est donc un endomorphisme orthogonal de E.*

Théorème. *Toute isométrie d'un espace affine euclidien est une application affine.*

Remarque. Une application affine $f : \mathcal{E} \to \mathcal{E}$ est une isométrie si et seulement si f transforme un (ou tout) repère orthonormé en repère orthonormé.

.

3.7. Définition. On appelle **déplacement** ou *isométrie directe* (respectivement, **antidéplacement** ou *isométrie indirecte*) de $\mathcal{E}$ toute isométrie f telle que $\text{dét}(df) = 1$ (respectivement, $\text{dét}(df) = -1$).

Les translations sont évidemment des déplacements. Les déplacements forment un sous-groupe dans $\mathrm{Iso}(\mathcal{E})$ d'indice 2 ; le produit de deux antidéplacements est un déplacement.

.

3.8. Définition. Soit $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$. La **projection orthogonale** sur $\mathcal{F}$ est la projection affine sur $\mathcal{F}$ parallèlement à $F^{\perp}$.

La **symétrie orthogonale** par rapport à $\mathcal{F}$ est la symétrie affine par rapport à $\mathcal{F}$ parallèlement à $F^{\perp}$. Une symétrie orthogonale est une isométrie ; elle est directe si et seulement si la codimension de $\mathcal{F}$ est paire (c'est à dire, $dim(\mathcal{E}) - dim(\mathcal{F})$ est paire).

On appelle **reflexion** une symétrie orthogonale par rapport à un hyperplan ; une reflexion est une isométrie indirecte (un antidéplacement).

Etant donné deux points A et B de $\mathcal{E}$, il existe une unique réflexion échangeant A et B : c'est la réflexion par rapport à l'hyperplan médiateur du segment $[AB]$.

Vu que toute isométrie vectorièle dans R^n s'écrit comme un produit d'au plus n réflexions, on a :

3.9. Proposition. Toute isométrie affine de $\mathcal{E}$ s'écrit comme un produit d'au plus $n+1$ réflexions ($n = \dim(\mathcal{E})$.)

.

3.10. Définition. Une transformation affine $f : \mathcal{E} \to \mathcal{E}$ est une **similitude** de rapport k si f multiplie toutes les distances par k :

$d(f(x), f(y)) = kd(x, y)$ pour tous $x, y \in \mathcal{E}$.

Une homothétie de rapport k est une similitude de rapport $\mid k \mid$.

Lemme. Une application affine $f : \mathcal{E} \to \mathcal{E}$ est une similtude de rapport k si et seulement si sa partie linéaire $df : E \to E$ est une similitude vectorielle : $\|df(v)\| = \pm k\|v\|$.

Remarque : les similitudes de $\mathcal{E}$ forment un groupe par rapport à la composition. Si f_i est une similitude de rapport k_i, $i = 1, 2$, alors $f_1 f_2$ est une similitude de rapport $k_1 k_2$.

3.11. Lemme. (i) Toute similitude de rapport $k \neq 1$ admet un point fixe (unique).

(ii) Toute similitude de rapport k est la composée d'une homothetie de rapport k et d'une isométrie.

.

3.12. Lemme. Soit T_v une translation de vecteur v et f une isométrie. Alors $fT_v = T_u f$ où $u = df(v)$. En particulier, T_v et f commutent si et seulement si $df(v) = v$.

.

3.13. Proposition.

Soit $f : \mathcal{E} \to \mathcal{E}$ une isométrie.

(i) Pour tout point $A \in \mathcal{E}$ il y a une unique décomposition $f = T_{\vec{a}} f_A$ où T est une translation de vecteur $\vec{a}$ et f_A une isométrie qui fixe A : $f_A(A) = A$.

(ii) Si df n'admet pas de vecteur fixe non-nul (donc si 1 n'est pas une valeur propre de df), alors f admet un unique point fixe.

.

Chosissons A comme origine, cela donne la bijection entre les points B et les vecteurs $v = \overrightarrow{AB}$. Soit $w = \overrightarrow{Af(B)}$. Alors $\vec{w} = \vec{a} + df(\vec{v})$, où $\vec{a} = \overrightarrow{Af(A)}$. Par rapport à cette vectorialisation f_A s'identifie avec l'application $\vec{v} \to df(\vec{v})$ et $T_{\vec{a}}$ avec $\vec{u} \to \vec{a} + \vec{u}$.

.

3.14 . Proposition. Il existe un point A tel que dans la décomposition $f = T_{\vec{a}} f_A$ on a $T_{\vec{a}} f_A = f_A T_{\vec{a}}$ (ce qui est équivalent à $df(\vec{a}) = \vec{a}$).

Classification des isométries en dimension 2 et 3.

La classification des isométries affines repose sur la classifcation des isométries vectorielles.

Lemme. Soit U une isométrie de l'espace vectoriel E.

1. dim $E = 2$. a) dét $U = 1$: U est une rotation.

b) dét $U = -1$: U est une réflexion.

2. dim $E = 3$. a) dét $U = 1$: U est une rotation autour d'un axe.

b) dét $U = -1$: U est une réflexion composée avec une rotation autour de l'axe orthogonal au plan de réflexion.

.

Isométries affines du plan

4.1. Proposition. (i) Tout déplacement du plan est une translation ou une rotation autour d'un point ; ici la rotation commute avec la translation.

(ii) Toute antidéplacement est le produit d'une réflexion par rapport à une droite avec une translation parallèle à cette droite (une "réflexion-translation") ; ici la reflexion commute avec la translation.

.

Isométries affines de l'espace.

4.2. Proposition. (i) Tout déplacement de l'espace est soit une translation soit le produit (commutatif) d'une rotation autour d'un axe et d'une translation parallèle à l'axe de rotation ("vissage").

(ii) Toute antidéplacement est soit le produit (commutatif) d'une réflexion par rapport à un plan avec une translation parallèle à ce plan ("réflexion-translation"), soit le produit (commutatif) d'une réflexion par rapport à un plan avec une rotation autour d'un axe orthogonal à ce plan ("réflexion-rotation").

.

Similitudes planes et nombres complèxes.

On identifie le plan euclidien R^2 avec le plan complèxe $\mathbf{C}$.

4.3. Proposition. Toute similitude directe de $\mathbf{C}$ s'écrit $s(z) = az+b$, avec $a \neq 0$. Toute similitude indirecte s'écrit $s(z) = a\bar{z} + b$, $a \neq 0$.

Le rapport d'une telle similitude est $\mid a \mid$.

En particulier, c'est une isométrie si et seulement si $\mid a \mid= 1$.