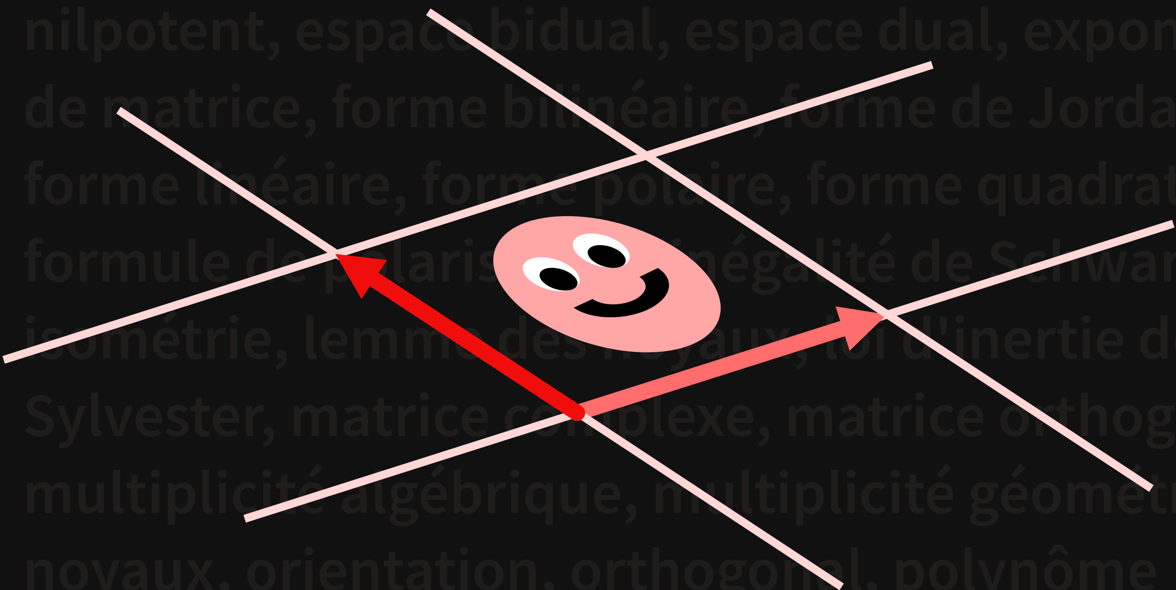
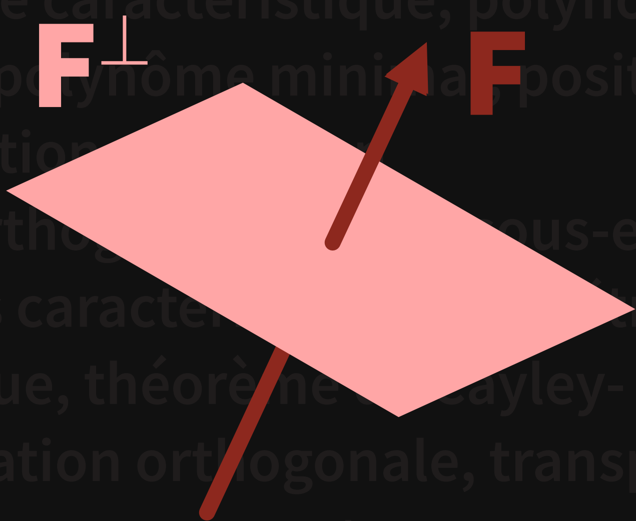
X31M020.

Algèbre Linéaire et Bilinéaire II.





Algèbre Linéaire et Bilinéaire II.

Ceci est un de 4 *pdf* de préparation à la L3. Il ne contient que les notions et connaissances qui seront enseignées l'année prochaine d'après le programme. Il n'y a pas de preuves; le but est juste de s'introduire aux concepts, pas de faire un cours entier (et rigoureux).

Les pages sont essentiellement (et en fait, quasiment entièrement) basées sur les polycopiés des années précédentes dans les matières ; il n'y a pas grand-chose ici, qu'on ne trouverait pas dans le cours.

Chaque page-fiche commence par une motivation; une sorte de dialogue entre un personnage D (Deuxième année), et un autre T (Troisième année), que vous pourrez appeler, dans votre tête, comme vous le souhaitez (*e.g.* <u>D</u>amien & <u>T</u>om Worro).

Sommaire.

Basé sur le polycopié (de Licence de maths, 3e année) *Résumé de cours X31M020 Algèbre linéaire et bilinéaire II*, Xue Ping WANG, 2020.

→ Rappels + polynômes.

Si on trouve comment appliquer un polynôme à des applications linéaires, on pourra faire le lien avec ses valeurs propres.

■ La motivation.

D – Toutes les symétries s doivent satisfaire $s^2 = Id$ (où s^2 veut dire $s \circ s$); pourrait-on dire que les symétries sont les racines de $P(X) = X^2 - Id$ appliqué aux applications linéaires ?
T – Oui, on dit que le polynôme $X^2 - 1$ annule les symétries ; et on remarque que les racines de P considéré comme un polynôme $(X^2 - 1)$ sur les réels sont +1 et -1, qui sont exactement les valeurs propres des symétries. Et pour une projection ?
D – On a $p^2 = p$, donc $X^2 - X$ annule les projections ; et on retrouve bien 0 et 1, valeurs propres des projections.
T – Eh oui ! Sauriez-vous montrer que si P annule f, alors toute valeur propre de f doit annuler P ?

■ Le cadre.

On se place dans E un e.v. sur $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) de dimension finie n, et on note $\mathcal{L}(E)$ l'espace des endomorphismes de E.

■ Les valeurs, vecteurs propres.

Pour $A \in \mathcal{M}_n(\mathbb{K})$, un scalaire λ est une valeur propre de A si il existe $x \in \mathbb{K}^n$ non-nul tel que $Ax = \lambda x$; et alors x est un vecteur propre. On pose $E_\lambda = \{x \in \mathbb{K}^n \mid Ax = \lambda x\}$, le sous-espace propre associé à λ. Et

$$\chi_A(\lambda) := \det(A - \lambda I_n) = (-1)^n \lambda^n + (-1)^{n-1} tr(A) \lambda^{n-1} + \ldots + \det(A)$$

est appelé polynôme caractéristique de A; ses racines sont exactement les valeurs propres de A. Pour λ_0 valeur propre, la multiplicité algébrique de λ_0 est l'ordre de la racine λ_0 de χ_A, autrement dit l'entier m_0 tel que $\chi_A(\lambda) = (\lambda - \lambda_0)^{m_0} q(\lambda)$ avec q polynôme ne s'annulant pas en λ_0. Et la multiplicité géométrique de λ_0 est la dimension de E_{λ_0}. Et : $\dim E_{\lambda_0} \leq m_0$.

■ La diagonalisation.

On dit que A est diagonalisable dans $\mathbb{K}$ si on peut trouver P inversible et D diagonale telle que $A = PDP^{-1}$. Voici des conditions :

• *ssi* χ_A est scindé dans $\mathbb{K}$, et pour toute valeur propre μ de A, les multiplicités algébrique et géométrique sont égales.
• *ssi* pour $\mu_1 \ldots \mu_k$ sont les valeurs propres de A, $E = E_{\mu 1} \oplus \ldots \oplus E_{\mu k}$.
• si A admet n valeurs propres distinctes.

Dans le cas de A matrice symétrique réelle, A est diagonalisable dans $\mathbb{R}$ par une matrice orthogonale Q.

■ La trigonalisation.

On dit que A est trigonalisable dans $\mathbb{K}$ si on peut trouver P inversible telle que $P^{-1}AP$ soit triangulaire supérieure. La condition la plus utile est : A est trigonalisable *ssi* χ_A est scindé dans $\mathbb{K}$. Ainsi, toute matrice carrée est trigonalisable dans $\mathbb{C}$.

■ Le polynôme d'un endomorphisme.

Pour $P(\lambda) = a_0 + a_1\lambda + \ldots + a_m\lambda^m$ un polynôme, et $f \in \mathcal{L}(E)$, on pose $f^k = f \circ \ldots \circ f$ (k fois) et $f^0 = I_E$, et posons $P(f) = a_0 I_E + a_1 f + \ldots + a_m f^m$, ainsi $P(f) \in \mathcal{L}(E)$; en particulier, on a $P_1(f) \circ P_2(f) = P_2(f) \circ P_1(f)$ pour P_1, P_2 deux polynômes de $\mathbb{K}[\lambda]$; si $f, g \in \mathcal{L}(E)$ commutent ($f \circ g = g \circ f$), alors $P_1(f) \circ P_2(g) = P_2(g) \circ P_1(f)$; et si λ_0 est valeur propre de f, alors $P(\lambda_0)$ est une valeur propre de P(f).

■ Le lemme des noyaux.

Si $P_1, P_2, \ldots P_k \in \mathbb{K}[X]$ sont des polynômes premiers entre eux et $P = P_1 P_2 \ldots P_k$, alors le lemme des noyaux énonce que :

$$\ker P(f) = \ker P_1(f) \oplus \ldots \oplus \ker P_k(f)$$

■ Les polynômes annulateurs.

Un polynôme annulateur de $f \in L(E)$ est un polynôme $P \in \mathbb{K}[X]$ tel que $P \neq 0$ et $P(f) = 0$. On sait qu'il doit en exister un, comme $\mathcal{L}(E)$ est de dimension n^2, et donc I_E, f, ... f^{n^2} sont liés. On a en particulier que tout polynôme annulateur a toutes les valeurs propres de f comme racines, mais ce ne sont pas obligatoirement les seules racines du polynôme !

■ Le théorème de Cayley-Hamilton.

Ce théorème a plusieurs formulations; la plus utile ici, est que le polynôme caractéristique de f, est un polynôme annulateur de f.

■ Le polynôme minimal.

On appelle polynôme minimal de f le polynôme annulateur de f, unitaire de plus petit degré; on le note $m_f(\lambda)$. Il divise tout polynôme annulateur de f. De plus, il est unique, et λ_0 est valeur propre de f *ssi* $m_f(\lambda_0) = 0$. Ainsi, f est diagonalisable *ssi* $m_f(\lambda)$ est scindé à racines simples sur K.

→ Sous-espaces caractéristiques.

Mais ... et si les multiplicités ne sont pas égales ?
Mieux vaut parler de sous-espaces caractéristiques.

La motivation.

D – Pour l'endomorphisme f de $\mathbb{R}^2$ défini par $(x,y)\mapsto(x-y,x+3y)$, on a le polynôme caractéristique $\chi(\lambda) = (\lambda-2)^2$, donc 2 a pour multiplicité algébrique 2 ; mais ker(f – 2id) = Vect((-1,1)) donne que 2 a pour multiplicité géométrique 1, et on ne peut pas diagonaliser ! On peut trigonaliser, mais comment fait-on ?
T – Posons g = f-2id ; alors $(-1,1)\in\ker g$, et $(1,0)\in\ker g^2$ (vu que $g^2 = 0$), avec g((1,0)) = (-1,1) ; donc dans la base {(-1,1), (1,0)}, g s'écrit comme $(u,v)\mapsto(v,0)$, et donc f = g+2id s'écrit comme $(u,v)\mapsto(2u+v,2v)$, qui est représentée par une matrice triangulaire. En général, c'est plus difficile; on doit s'appuyer sur les sous-espaces caractéristiques, généralisation des sous-espaces propres.

Les sous-espaces caractéristiques.

Pour f trigonalisable, le polynôme caractéristique s'écrit :

$$\chi_f(\lambda) = (-1)^n(\lambda-\lambda_1)^{\alpha_1}\dots(\lambda-\lambda_p)^{\alpha_p}$$

où $\lambda_i\neq\lambda_j$ si $i\neq j$. On appelle $N_{\lambda j} = \ker[(f-\lambda_j I)^{\alpha_j}]$ le sous-espace caractéristique de f associé à λ_j. Et $\ker[(f-\lambda_j I)^k] \subseteq \ker[(f-\lambda_j I)^m]$ si k<m. Le polynôme minimal de f s'écrit $(-1)^n(\lambda-\lambda_1)^{m_1}\dots(\lambda-\lambda_p)^{m_p}$ (avec $1\leq m_i\leq\alpha_i$) et on a, théorème :
- $E = N_{\lambda 1}\oplus\dots\oplus N_{\lambda p}$;
- $N_{\lambda j} = \ker[(f-\lambda_j I)^{m_j}]$;
- $\dim N_{\lambda j} = \alpha_j$.

Et $N_{\lambda j}$ est stable par f, et la restriction g_j de $f-\lambda_j I$ sur $N_{\lambda j}$ est nilpotente, d'indice de nilpotence m_j (voir plus bas).

La restriction.

Si $f\in L(E)$ est diagonalisable (resp. trigonalisable), et $F\subseteq E$ s.e.v. de E, alors $g = f_{|F}$ (f restreint à F) est aussi diagonalisable (resp. trigonalisable).

Les endomorphismes nilpotents.

Un endomorphisme g est nilpotent s'il existe k entier, tel que $g^k = 0$. Le plus petit m tel que $g^m = 0$ est l'indice de nilpotence de g. Si g est nilpotent d'indice de nilpotence m, alors $\chi_g(\lambda) = (-1)^n\lambda^n$, $m_g(\lambda) = \lambda^m$. En plus, on a la suite d'inclusions strictes $\ker g \subsetneq \ker g^2 \subsetneq \dots \subsetneq \ker g^m = E$.

La forme de Jordan, cas nilpotent.

Un bloc ou matrice de Jordan (d'ordre d) est une matrice de la forme

$$J_d(\mu) = \begin{pmatrix} \mu & 1 & 0 & \cdots & 0 \\ 0 & \mu & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \mu & 1 \\ 0 & \cdots & \cdots & 0 & \mu \end{pmatrix} \in \mathcal{M}_d(\mathbb{K})$$

On notera J_d pour $J_d(0)$. Pour tout endomorphisme $g\in\mathcal{L}(E)$ nilpotent (d'indice de nilpotence m), il existe une base e de E telle qu'on puisse écrire g comme :

$$[g]_e = \begin{pmatrix} J_{d_1} & 0 & \cdots & \cdots & 0 \\ 0 & J_{d_2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & J_{d_k} \end{pmatrix}$$

où $m=d_1 \geq \dots \geq d_k \geq 1$ et $d_1+\dots+d_k = n$.

La forme de Jordan, cas général.

Il y a bien sûr un cas général, plus compliqué. Pour f trigonalisable, la restriction g_j de $f-\lambda_j I$ sur $N_{\lambda j}$ est nilpotente pour tout λ_j, donc il existe une base de $N_{\lambda j}$ telle que f restreint à $N_{\lambda j}$ s'écrive comme une diagonale de λ_j, avec des 1 ou des 0 sur la diagonale juste au-dessus, et des 0 partout ailleurs; ce qu'on peut écrire

$$[f|_{N_{\lambda_j}}]_{e^{(j)}} = \mathrm{Diag}(J_{d_{1,j}}(\lambda_j), \cdots, J_{d_{k_j,j}}(\lambda_j))$$

où $d_{1,j} = m_j$, et $d_{1,j}+\dots+d_{kj,j} = \alpha_j$; en mettant ensemble ces blocs pour chaque valeur propre λ_j selon la diagonale d'une matrice, on obtient une matrice semblable à celle de départ (donc qu'on peut obtenir par un changement de base bien choisi), mais triangulaire.

La décomposition D+N.

La décomposition D+N, aussi dite de Dunford, elle permet d'écrire tout A trigonalisable (de manière unique) comme D+N, où DN = ND, D est diagonalisable et N nilpotente. Cette décomposition permet, pour une matrice n×n donnée, d'en expliciter une puissance quelconque, et d'en expliciter l'exponentielle.

Les applications.

En écrivant A = D+N avec DN = ND et $N^{m+1} = O$ (avec $N^m \neq O$). Alors, on peut appliquer la formule du binôme vu que D et N commutent; et

$$A^q = \sum_{j=0}^{m} \binom{q}{j} D^{q-j} N^j$$

(quand q≥m). On peut ainsi donc écrire l'exponentielle d'une matrice e^{tA}, en utilisant que $e^{tA} = e^{tD}e^{tN}$ comme D et N commutent, et donc,

$$e^{tA} = P \begin{pmatrix} e^{tD'_{1,1}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{tD'_{n,n}} \end{pmatrix} P^{-1} \cdot \left(\sum_{j=0}^{m} \frac{t^j}{j!} N^j\right)$$

Où $D = PD'P^{-1}$ avec D diagonale. Cela permet de donner une solution aux systèmes différentiels à coefficients constants de la forme $x'(t) = Ax(t)$ avec $x(0) = x_0$ fixé pour A une matrice de $\mathcal{M}_n(\mathbb{K})$; cette solution est $x(t) = e^{tA}x_0$.

→ Dualité.

Pour tout espace vectoriel, on a une sorte d'espace miroir, qui fonctionne à l'envers.

■ La motivation.

D – En physique, on utilise souvent de formes linéaires ; pourquoi ?
T – Par exemple, on peut construire des objets mathématiques qui sont linéaires en plusieurs dimensions, les tenseurs, à partir de vecteurs et formes linéaires. En particulier, les formes bilinéaires peuvent être vues comme de tenseurs formés avec des produits de formes linéaires. Les tenseurs servent en ingénierie (e.g. tenseur des contraintes), mais plus notablement en relativité générale ; des champs de tenseur (fonctions qui associent un tenseur à chaque endroit de l'espace) permettent de décrire la répartition de matière-énergie, et la courbure. Et ça sert aussi en physique quantique, dans la notation bra-ket.

■ Les formes linéaires.

Une forme linéaire sur E, un espace vectoriel, est une application linéaire de E dans $\mathbb{K}$. L'ensemble des formes linéaires, $\mathcal{L}(E,\mathbb{K})$, est un espace vectoriel, l'espace dual de E, qu'on note ainsi E*. Notons qu'on peut écrire toute forme linéaire comme $f(x) = a_1x_1+\ldots+a_nx_n$ (où $x = (x_1,\ldots,x_n)$), avec un unique choix de a_1, … et a_n. On peut ainsi trouver une base $e^* = (e^*_1, \ldots e^*_n)$ de E* pour toute base $e = (e_1, \ldots e_n)$ de E, telle que $e^*_j(e_k) = \delta_{jk}$; alors e* est la base duale de e (et e la base préduale de e*). Pour i, e deux bases de E et P la matrice de passage de i à e, alors la matrice de passage P de la base duale i* à e* est donnée par $P^* = {}^tP^{-1}$.

■ L'orthogonalité.

Pour un sous-ensemble $F\subseteq E$, on peut définir son orthogonal F° comme $F° = \{f\in E^* : f(x) = 0, \forall x\in F\}$; qui est donc un s.e.v. de E*; et réciproquement, l'orthogonal G° de $G\subseteq E^*$ est défini comme $G° = \{x\in E : f(x) = 0, \forall f\in G\}$. Si F est un s.e.v. de E, alors dim E = dim F + dim F°, et (F°)° = F.

■ Le bidual.

Posons $E^{**} = \mathcal{L}(E^*,\mathbb{K})$ l'espace bidual de E. Par exemple, les applications $i_x:E^*\to\mathbb{K}$ définies par $i_x(f) = f(x)$ pour tout $f\in E^*$ sont des éléments de E**. D'ailleurs : l'application $i:E\to E^{**}$ définie par $i(x) = i_x$ est un isomorphisme de E sur E** (l'injection canonique de E dans E**).

■ La transposée.

Soit $A\in\mathcal{L}(E,F)$; on définit alors sa transposée ${}^tA$ comme une application $F^*\to E^*$, avec pour tout $f\in F^*$, ${}^tAf = g$ où g est la forme linéaire satisfaisant g(x) = f(Ax) (pour tout $x\in E$). On a, pour $A,B\in\mathcal{L}(E)$, et $\lambda,\mu\in K$: ${}^t\mathrm{Id}_E = \mathrm{Id}_{E^*}$, ${}^t(\lambda A+\mu B) = \lambda{}^tA+\mu{}^tB$, et ${}^t(A\circ B) = {}^tB\circ{}^tA$. Au niveau matriciel, on a $[{}^tA]_{e^*} = {}^t[A]_e$, où e est une base E, et e* sa base duale dans E*; et ${}^t[A]_e$ est la transposée de la matrice $[A]_e$. Et : rang A = rang ${}^tA$.

→ Formes quadratiques.

On peut étudier les formes bilinéaires, mais la signature nous dit qu'il y en a (n+1)(n+2)/2, à changement de base près.

La motivation.

D – Si j'ai une équation quadratique, comment faire pour décrire la forme de la surface ?
T – Il y a toujours une base qui permet d'écrire une forme bilinéaire donnée, comme une somme/différences de carrés ; et alors, la forme dépend uniquement de la signature. De là, il s'agit juste de savoir à quoi correspond une signature donnée ; ou bien on l'apprend, ou bien on le devine.
D – Et la base trouvée, est-ce qu'elle peut servir à quelque chose ?
T – Oui ; par exemple en statistique, ce serait les « composantes principales » (méthode *PCA* en anglais).

Les formes bilinéaires.

Une application ϕ:E×E→$\mathbb{K}$ est dite forme bilinéaire si pour tout x fixé, y→ϕ(x,y) est linéaire, et si pour tout y fixé, x→ϕ(x,y) est linéaire. Elle est dite symétrique si ϕ(x,y) = ϕ(y,x) pour tous x,y∈E. Pour une base e = {e_1, … e_n} de E, on définit la matrice de ϕ par $[\phi]_e = (\phi(e_i,e_j))_{1\le i,j\le n}$; et alors

$$\phi(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{n}\phi(e_i,e_j)x_iy_j =^t [x]_e\cdot[q]_e\cdot[y]_e$$

ϕ est symétrique *ssi* $[\phi]_e$ est symétrique; et si P est la matrice de passage de e à une autre base e', alors $[\phi]_{e'}$ = ${}^tP[\phi]_eP$ (et rang $[\phi]_e$ = rang $[\phi]_{e'}$; on pose donc rang ϕ = rang $[\phi]_e$); on dit que A et B matrices carrées sont congruentes s'il existe P matrice inversible telle que A = ${}^tPBP$.

La positivité.

Une forme bilinéaire symétrique ϕ est dite positive (noté ϕ≥0) si ϕ(x,x)≥0 pour tout x, et définie positive, si ϕ≥0 et si ϕ(x,x) = 0 implique x = 0.

Les formes quadratiques.

Une application q:E→K qui peut s'écrire q(x) = ϕ(x,x) avec ϕ forme bilinéaire symétrique est dite forme quadratique; ϕ en est sa forme polaire. Et q est dite positive si ϕ l'est; on pose $[q]_e = [\phi]_e$ et rang q = rang ϕ. On a quelques propriétés intéressantes :
- q(λx) = λ²q(x) pour tout x∈E, λ∈K;
- q(x+y) = q(x)+q(y)+2ϕ(x,y) et q(x+y)+q(x-y) = 2(q(x)+q(y));
- et les formules de polarisation, qui donnent ϕ en fonction de q : ϕ(x,y) = ½(q(x+y)–q(x)–q(y)) = ¼(q(x+y)–q(x-y)).

Les noyaux.

Posons G = {x; ϕ(x,y) = 0, ∀y∈E} (appelé noyau à gauche), et D = {y; ϕ(x,y) = 0, ∀x∈E} (noyau à droite). Si ϕ est bilinéaire symétrique, G = D, et on pose ker ϕ = G = D. Et pour le q associé, on pose ker q = ker ϕ. Mais attention ! En général, Iso(q) = {x; q(x) = 0} est différent de ker q. C'est le cas si q est une forme quadratique positive sur E. On a dim G = dim D = dim E - rang ϕ, pour tout ϕ. On dit que q et ϕ sont non-dégénérées si rang q = dim E, donc si ker q = {0}.

Les inégalités.

- celle de Schwarz : si ϕ est une forme bilinéaire symétrique positive sur E, |ϕ(x,y)|²≤ϕ(x,x)·ϕ(y,y).

- celle de Minkowski : √q(x+y) ≤ √q(x) + √q(y).

La décomposition en somme de carrés.

Le théorème important : pour q une forme quadratique sur E, il existe un entier r, r scalaires non-nul λ_1, … λ_r et r formes linéaires linéairement indépendantes L_1, … L_r sur E, tels qu'on puisse écrire q(x) comme $\Sigma_{1\le j\le r}\lambda_j(L_j(x))^2$, et r = rang q. Pour montrer ça constructivement, on utilise, après l'avoir défini, l'algorithme de Gauss (*cf* la page Méthodes).

Les bases orthogonales.

Pour q une forme quadratique, deux vecteurs x,y de E sont orthogonaux quand ϕ(x,y) = 0; et donc une base e = (e_1, … e_n) de E est dite orthogonale relativement à q si tous les vecteurs sont orthogonaux deux-à-deux; c'est-à-dire si $\phi(e_i,e_j)$ = 0 pour tous i≠j. De par la décomposition en carrés, pour n'importe quelle forme quadratique, il existe une base qui est orthogonale relativement à elle. On définit l'orthogonal de F⊆E s.e.v. relativement à une forme quadratique q par $F_q^\perp$ = {x∈E : ϕ(x,y) = 0, ∀y∈F}. On a donc dim $F_q^\perp$ = dim E - dim F + dim (F∩ker q). Dans le cas q est non-dégénérée, dim $F_q^\perp$ = dim E - dim F.

La loi d'inertie de Sylvester.

Pour une forme quadratique réelle q sur E, il existe une base e = (e_1, … e_n) de E et deux entiers s et t tels que

$$q(x_1e_1+\cdots+x_ne_n) = \sum_{j=1}^{s}x_j^2 - \sum_{j=s+1}^{s+t}x_j^2$$

L'essentiel de la loi dit que le couple (s,t) est indépendant de la base choisie; c'est la signature de q; on note sgn(q) = (s,t). Corollaire : deux matrices réelles symétriques sont congruentes dans $\mathcal{M}_n(\mathbb{R})$ *ssi* elles ont la même signature.

Les transformations orthogonales.

Pour E un e.v. sur $\mathbb{K}$, et q:E→$\mathbb{K}$ quadratique non-dégénérée, on appelle transformation orthogonale relativement à q tout f∈L(E) telle que q(f(x)) = q(x) pour tout x∈E; on notera O(E,q) l'ensemble des transformations orthogonales relativement à q. Et "f est une transformation orthogonale relativement à q" est équivalent à "ϕ(f(x),f(y)) = ϕ(x,y) pour tous x,y∈E"; de plus, si f,g∈O(E,q), f∘g∈O(E,q); et si f∈O(E,q), alors f est inversible et f^{-1}∈O(E,q).

→ Compléments (1).

Des applications, des généralisations, et notamment le cas des isométries en trois dimensions.

La motivation.

D – Quand on lance un objet indéformable, comment peut-on le décrire mathématiquement ?
T – En choisissant un point particulier (e.g. centre de gravité), il suffit de décrire la trajectoire de ce point, et l'isométrie (de déterminant +1) que subit l'objet. Parce que l'objet ne garde pas forcément la même orientation ; mais comme il est indéformable, la distance entre deux points de l'objet ne change pas ; donc la transformation doit préserver la norme.
D – Et en 3d, les isométries de déterminant +1 sont des rotations, n'est-ce pas ?
T – Effectivement ; et on peut décrire une rotation 3d par une direction, et un angle. Mais ce n'est pas la seule façon …

Les produits scalaires.

Un produit scalaire sur E e.v. réel est une forme bilinéaire symétrique positive et non-dégénérée qu'on note $\langle\cdot,\cdot\rangle$; alors $\|x\| = \sqrt{q(x)}$ (où $q(x) = \langle x,x\rangle$) est une norme sur E. Et on appelle (E, $\langle\cdot,\cdot\rangle$) un espace vectoriel euclidien. Les (in)égalités qu'on a vues deviennent $\|x+y\|^2+\|x-y\|^2 = 2(\|x\|^2+\|y\|^2)$ et $\langle x,y\rangle = ½(\|x+y\|^2-\|x\|^2-\|y\|^2) = ¼(\|x+y\|^2-\|x-y\|^2)$; l'inégalité de Schwarz devient $|\langle x,y\rangle| \leq \|x\|\cdot\|y\|$. On peut trouver une base orthonormale e = (e_1, …, e_n) ; c'est-à-dire pour laquelle on a $\langle e_i, e_j\rangle = \delta_{ij}$, donc telle que $\|x\|^2 = \Sigma_j\, x_j^2$ quand $x = x_1e_1+\ldots+x_ne_n$.

Les formes linéaires (2).

Toute forme linéaire ℓ de E peut s'écrire, de manière unique, comme $\ell(x) = \langle a,x\rangle$ pour tout $x\in E$, pour un certain a.

L'orthogonal.

On écrit $x\perp y$ si $\langle x,y\rangle = 0$; on définit l'orthogonal de F par $F^\perp = \{x\in E;\ \langle x,y\rangle=0,\ \forall y\in F\}$. Alors on a, pour tout s.e.v. de E de dimension k, $E = F\oplus F^\perp$ et E admet une b.o.n. e = {e_1, … e_n} tq {e_1, … e_k} soit b.o.n. de F, et {e_{k+1}, … e_n} soit b.o.n. de $F^\perp$. Et : $u\in\mathcal{L}(E)$ est une projection orthogonale (sur un F, parallèlement à $F^\perp$) *ssi* $u^2 = u$ et Im $u\perp$ker u; u est alors la projection orthogonale sur Im u. Et pour tout b.o.n. {b_1, … b_k} de F, la projection orthogonale π_F de E sur F est donnée par

$$\pi_F(x) = \sum_{j=1}^{k}\langle b_j, x\rangle b_j$$

L'endomorphisme adjoint.

Pour $f\in L(E)$, on a que : pour chaque x, l'application $y\to\langle f(y),x\rangle$ est une forme linéaire, donc on peut l'écrire comme $\langle f(y),x\rangle = \langle y,h\rangle$ pour un certain h, et l'application g qui associe à x, le vecteur h correspondant, est appelé endomorphisme adjoint de f, noté f*. On a que $\langle f(x),y\rangle = \langle x,f^*(y)\rangle$ pour tout $x,y\in E$; et $[f^*]_e = {}^t[f]_e$; rang f = rang f*, $\chi_f(\lambda) = \chi_{f^*}(\lambda)$, $(f^*)^* = f$, Ker f = (Im f*)$^\perp$ et Im f = (Ker f*)$^\perp$; et on a $(\lambda f+\mu g)^* = \lambda f^*+\mu g^*$, et $(f\circ g)^* = g^*\circ f^*$.

Les endomorphismes auto-adjoints.

Ce sont les $f\in L(E)$ tels que f* = f, donc que $\langle f(x),y\rangle = \langle x,f(y)\rangle$. Si e est une b.o.n. de E, f est autoadjoint *ssi* $[f]_e$ est une matrice réelle symétrique. Et, si f* = f, alors il existe une b.o.n. de E formée de vecteurs propres de f. Ainsi, toute matrice réelle symétrique est diagonalisable par une matrice orthogonale.

Les matrices complexes.

Une matrice $A\in\mathcal{M}_n(\mathbb{C})$ est normale si $A\circ A^* = A^*\circ A$ où $A^* = {}^t\bar{A}$, et U est unitaire si $U\circ U^* = U^*\circ U = I_n$; et donc, toute matrice normale est diagonalisable par une matrice unitaire dans $\mathbb{C}$.

Le résultat.

Pour toute forme quadratique q dans un EVE (E, $\langle\cdot,\cdot\rangle$), il existe une b.o.n. e de E, et des λ_j tels que

$$q(x) = \sum_{j=1}^{n}\lambda_j x_n^2$$

Pour tout $x = x_1e_1+\ldots+x_ne_n\in E$.

Les isométries vectorielles.

Pour E un EVE, on note O(E) = O(E, $\|\cdot\|^2$) le groupes des isométries vectorielles relativement à $\|\cdot\|^2$ (c'est-à-dire q, ou $x\mapsto\langle x,x\rangle$). On a l'équivalence entre :

i. f est une isométrie vectorielle, i.e. préserve la norme;
ii. f conserve le produit scalaire : $\langle f(x),f(y)\rangle = \langle x,y\rangle$;
iii. la matrice de f dans une b.o.n. est une matrice orthogonale;
iv. f transforme toute base orthonormale de E, en une base orthonormale;
v. plus simplement : $f^{-1} = f^*$.

Notations : $O^+(E) = \{f\in O(E);\ \det f = +1\}$ et $O^-(E) = \{f\in O(E);\ \det f = -1\}$ contiennent les isométries directes et indirectes respectivement. $O(n) = \{A\in\mathcal{M}_n(\mathbb{R}),\ {}^tAA = I_n\}$ (matrice orthogonale) et $SO(n) = \{A\in O(n),\ \det A = +1\}$ sont des sous-groupes de $\mathcal{GL}_n(\mathbb{R})$.

→ Compléments (2).

Des applications, des généralisations, et notamment le cas des isométries en trois dimensions.

La motivation.

D – Quand on lance un objet indéformable, comment peut-on le décrire mathématiquement ?
T – En choisissant un point particulier (e.g. centre de gravité), il suffit de décrire la trajectoire de ce point, et l'isométrie (de déterminant +1) que subit l'objet. Parce que l'objet ne garde pas forcément la même orientation ; mais comme il est indéformable, la distance entre deux points de l'objet ne change pas ; donc la transformation doit préserver la norme.
D – Et en 3d, les isométries de déterminant +1 sont des rotations, n'est-ce pas ?
T – Effectivement ; et on peut décrire une rotation 3d par une direction, et un angle. Mais ce n'est pas la seule façon …

L'orientation.

Ici, le choix d'une base i d'un espace vectoriel réel E donne une orientation de celui-ci; une autre base e est alors dite directe (ou positive) si le déterminant de la matrice de passage de i à e est positif; et indirecte (ou négative) si ce déterminant est négatif.

Le cas de la dimension 3.

Soit $f\in O(E_3)$, et $A = [f]_i$ (pour i la base qui oriente E_3, un EVE de dimension 3); alors si $\gamma = \det f$, γ est valeur propre de f, et il existe $\theta\in[0,2\pi[$ tel que $[f]_e$ puisse s'écrire, dans une base euclidienne e bien choisie, comme :

$$[f]_e = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$$

En particulier si det f = 1, f est une rotation autour de $\mathrm{Vect}(e_3)$; si det f = -1, f est une rotation (autour de $\mathrm{Vect}(e_3)$) suivi d'une réflexion par rapport au plan $\mathrm{Vect}(e_1,e_2)$.

Les rotations.

Pour u, v deux vecteurs libres dans E_3, une orientation du plan P = Vect(u,v) est le choix d'un vecteur n normal au plan; on peut alors calculer l'angle orienté $\theta = (u,v)_n$, comme solution de $\cos\theta = \langle u,v\rangle/\|u\|\cdot\|v\|$; et $\sin\theta = \det([u]_i,[v]_i,[n]_i)/\|u\|\cdot\|v\|\cdot\|n\|$. Pour simplifier, on notera $\det([u]_i,[v]_i,[n]_i)$, appelé produit mixte, par [u,v,n]. Voici donc comment calculer le θ d'une isométrie, si on connaît la base (e_1,e_2,e_3) qui permet de l'écrire comme plus haut. On sait que θ doit satisfaire $\cos\theta = (\mathrm{Tr}\, f - \gamma)/2$ et $\sin\theta = [x,f(x),n]/\|x\|^2\cdot\|n\|$, où on prend $n = e_3$ et x un vecteur non-nul du plan $\mathrm{Vect}(e_1,e_2)$.

La réduction d'endomorphismes orthogonaux.

Soit $u\in O(E)$. Alors on peut trouver une base orthogonale e et $p,q,r\in\mathbb{N}$ tels que $[u]_e = \mathrm{Diag}(I_p, -I_q, R_1, \dots R_r)$ avec

$$R_j = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$$

avec chaque $\theta_j\in[0,2\pi[$ (et $\theta_j \neq \pi$). En particulier, pour une matrice $A\in O(n)$, on peut trouver une matrice Q orthogonale telle que ${}^tQAQ$ soit sous la forme décrite plus haut. On peut même dire que u est un produit commutatif de q réflexions et r rotations, avec p = dim ker(u-Id), q = dim ker(u+Id), et 2r = n-p-q.

Les symétries vectorielles.

Une symétrie orthogonale de E par rapport à F s.e.v. de E est l'application linéaire s_F qui, à tout x = y+z décomposé dans $F\oplus F^\perp$, associe y-z. Autrement dit, s_F est linéaire, et $s_F(y) = y$ si $y\in F$, et $s_F(z) = -z$ si $z\in F^\perp$. On a que $s_F = 2\pi_F - \mathrm{Id} = \pi_F - \pi_{F^\perp}$; et $s_F = -s_{F^\perp}$. Et un endomorphisme $u\in\mathcal{L}(E)$ est une symétrie orthogonale *ssi* $u^2 = \mathrm{Id}$, et $\ker(u-\mathrm{Id})\perp\ker(u+\mathrm{Id})$. En particulier, réflexion désigne une symétrie orthogonale par rapport à un hyperplan, et retournement (demi-tour) une symétrie orthogonale par rapport à une droite. De manière intéressante, le groupe O(E) en engendré par l'ensemble des réflexions; en particulier, tout $u\in O(E)$ est le produit de n-p réflexions, où p = dim ker(u-Id).

→ Méthodes.

Voici des méthodes pour faire certains calculs en Algèbre Linéaire et Bilinéaire ; certaines servent de preuves.

Le polynôme minimal.

$\chi_f(\lambda)$ se décompose en $p_1(\lambda)^{\alpha_1}\dots p_k(\lambda)^{\alpha_k}$ avec les p_j des polynômes irréductibles premiers entre eux (et $\alpha_k \geq 1$), il faut chercher le polynôme minimal de la forme $m_f(\lambda) = p_1(\lambda)^{m_1}\dots p_k(\lambda)^{m_k}$ avec $1\leq m_j \leq \alpha_j$ pour tout j. En particulier, si tous les p_j sont de degré 1 (donc χ_f scindé), on peut calculer $k_j = \dim E^{\lambda_j}$; alors on sait que si $k_j = 1$, $m_j = \alpha_j$; si $k_j = \alpha_j$, $m_j = 1$; et si, plus généralement, $2\leq k_j \leq \alpha_{j-1}$, alors $2\leq m_j \leq \alpha_j - k_j + 1$.

La réduction de Jordan d'un endomorphisme nilpotent.

Soit $g\in\mathcal{L}(E)$ (E de dim n) nilpotent. D'abord, déterminer m l'indice de nilpotence en calculant g^i jusqu'à obtenir 0. Puis, calculer $E_j = \ker g_j$ pour j de 1 à m-1. Choisissons $b_m\in E_m \backslash E_{m-1}$; on pose alors $b_{m-1} = g(b_m)$, $b_{m-2} = g(b_{m-1})$, … $b_1 = g(b_2)$. On choisit ensuite des s.e.v. F_m, … F_1 tels que $E = E_{m-1}\oplus F_m$, $E_{m-1} = E_{m-2}\oplus g(F_m)\oplus F_{m-1}$, … $E_1 = g^{m-1}(F_m)\oplus\dots\oplus g(F_2)\oplus F_1$; il est possible que certains F_j soient nuls !

La décomposition D+N (1).

Soit $A\in\mathcal{M}_n(\mathbb{K})$ trigonalisable, et f l'endomorphisme de $\mathbb{K}^n$ correspondant ($A = [f]_i$, avec i base canonique). On calcule pour chaque sous-espace caractéristique $N_{\lambda j}$ une base $\{e_1^{(j)}, \dots e_{\alpha j}^{(j)}\}$; en les concaténant, on obtient une base e de $\mathbb{K}^n$. Ensuite, on pose la matrice $\mathrm{Diag}(\lambda_1 I_{\alpha 1}, \dots \lambda_p I_{\alpha p})$; elle représente, dans la base e, un certain endomorphisme d (tel que $de_i^{(j)} = \lambda_j e_i^{(j)}$). Maintenant, avec P la matrice de passage de i à e, on pose $D = P[d]_e P^{-1}$; alors, avec N = A-D, on a une décomposition D+N (D diagonalisable, N nilpotente, A=D+N, DN = ND).

La décomposition D+N (2).

On trigonalise A, c'est-à-dire on trouve une matrice de passage P telle que $P^{-1}AP = B$, avec B triangulaire. Ensuite, on pose D' la diagonale de B (et N' = B-D'). Et puis, avec $D = PDP^{-1}$, et N = A–D, on trouve une décomposition D+N.

Les noyaux G et D.

Pour e base de E, et $A = [\phi]_e$, D s'obtient comme

$$D = \{x = x_1e_1 + \dots + x_ne_n : A\begin{pmatrix}x_1\\ \vdots \\ x_n\end{pmatrix} = 0\}$$

Et similairement, G est donné par

$$G = \{x = x_1e_1 + \dots + x_ne_n : A^t\begin{pmatrix}x_1\\ \vdots \\ x_n\end{pmatrix} = 0\}$$

L'algorithme de Gauss.

On peut écrire toute forme quadratique comme un polynôme quadratique en n variables :

$$q(x_1e_1 + \dots + x_ne_n) = \sum_{j=1}^{n} a_{jj}x_j^2 + 2\sum_{1\leq i<j\leq n} a_{ij}x_ix_j$$

Alors :

1. Si il existe un aii non-nul, alors en posant

$$L(x) = 1/2a_{ii}\,(\partial q/\partial x_i)(x)$$

$q_1(x) = q(x) - a_{ii}(L(x))^2$ contient au moins une variable en moins; et donc on recommence dessus.

2. Si $a_{ii} = 0$ pour tous i, alors il y a un a_{ij} non-nul; alors en posant $y_i = (x_i+x_j)/2$, $y_j = (x_i-x_j)/2$, et $y_k = x_k$ pour les autres k, on peut écrire q en fonction des y_1 … y_n; et alors, parmi les nouveaux coefficients $\tilde{a}_{ij}$, on a $\tilde{a}_{ii} = 2a_{ij}$ est non-nul, on peut maintenant appliquer l'étape 1.

En répétant, on diminue le nombre de variables, jusqu'à ce qu'il n'y en ait plus.

Le procédé d'orthonormalisation de Gram-Schmidt.

Soit donc $\{e_1, \dots e_k\}$ base d'un e.v. F (par exemple, F s.e.v. d'un e.v. E), muni d'un produit scalaire $\langle\cdot,\cdot\rangle$; on pose $\|u\| = \sqrt{\langle u,u\rangle}$. Alors on peut construire une base orthonormale de F en posant : $q_1 = e_1/\|e_1\|$, et puis récursivement

$$\tilde{e}_j = e_j - \sum_{l=1}^{j-1}\langle e_j, q_l\rangle q_l$$

et $q_j = \tilde{e}_j/\|\tilde{e}_j\|$. Alors non-seulement $\{q_1, \dots q_k\}$ est orthonormale, mais en plus, on a que $\mathrm{Vect}(q_1, \dots q_j) = \mathrm{Vect}(e_1, \dots e_j)$ pour tout j de 1 à k.