Université Claude Bernard–Lyon I
CAPES de Mathématiques : Algèbre linéaire
Année 2006–2007

Chapitre 0
Vecteurs

1° La structure

a) Groupe

On appelle *groupe* un ensemble $\mathbb{G}$ muni d'une opération

$$\begin{array}{rccc} *: & \mathbb{G}\times\mathbb{G} & \longrightarrow & \mathbb{G} \\ & (\lambda,\mu) & \longmapsto & \lambda*\mu \end{array}$$

satisfaisant les conditions suivantes :

- Associativité : pour tous $\lambda, \mu, \nu \in \mathbb{G}$, on a : $(\lambda*\mu)*\nu = \lambda*(\mu*\nu)$; on note alors le résultat $\lambda*\mu*\nu$, sans ambiguïté ;
- Neutre : il existe $e \in \mathbb{G}$ tel que pour tout $\lambda \in \mathbb{G}$, $e*\lambda = \lambda*e = \lambda$; cet élément est alors unique ; si $*$ est notée $+$, on note $e = 0$; si $*$ est un produit $\cdot$, on note $e = 1$;
- Opposé/inverse : pour tout $\lambda \in \mathbb{G}$, il existe $\lambda' \in \mathbb{G}$ tel que $\lambda*\lambda' = \lambda'*\lambda = e$; alors λ' est unique ; si $* = +$, on note $\lambda' = -\lambda$; si $* = \cdot$, on note $\lambda' = \lambda^{-1}$.

On dit de plus que $\mathbb{G}$ est *commutatif* si pour tous $\lambda, \mu \in \mathbb{G}$, on a : $\lambda*\mu = \mu*\lambda$.

b) Corps

On appelle *corps* un ensemble $\mathbb{K}$ muni de deux opérations (addition et produit)

$$\begin{array}{rccc} +: & \mathbb{K}\times\mathbb{K} & \longrightarrow & \mathbb{K} \\ & (\lambda,\mu) & \longmapsto & \lambda+\mu \end{array} \quad \text{et} \quad \begin{array}{rccc} \cdot: & \mathbb{K}\times\mathbb{K} & \longrightarrow & \mathbb{K} \\ & (\lambda,\mu) & \longmapsto & \lambda\,\mu \end{array}$$

telles que :

- $\mathbb{K}$ soit un groupe commutatif pour la loi $+$; on note 0 son neutre ;
- $\mathbb{K}\setminus\{0\}$ soit un groupe commutatif pour la loi $\cdot$; on note 1 son neutre ; en particulier :
  - $\forall\lambda \in E, \quad 1\cdot\lambda = \lambda$;
  - $\forall\lambda,\mu,\nu \in \mathbb{K}, \quad (\lambda\cdot\mu)\cdot\nu = \lambda\cdot(\mu\cdot\nu)$.
- les lois $+$ et $\cdot$ soient compatibles, c'est-à-dire :
  - $\forall\lambda,\mu,\nu \in \mathbb{K}, \quad (\lambda+\mu)\cdot\nu = \lambda\cdot\nu+\mu\cdot\nu$;
  - $\forall\lambda,\mu,\nu \in \mathbb{K}, \quad \lambda\cdot(\mu+\nu) = \lambda\cdot\mu+\lambda\cdot\nu$;

c) Espace vectoriel

Soit $\mathbb{K}$ un corps. On appelle *espace vectoriel sur* $\mathbb{K}$ un ensemble E muni de deux loi (addition des vecteurs et produit par un scalaire)

$$\begin{array}{rccc} +: & E\times E & \longrightarrow & E \\ & (v,w) & \longmapsto & v+w \end{array} \quad \text{et} \quad \begin{array}{rccc} \cdot: & \mathbb{K}\times E & \longrightarrow & E \\ & (\lambda,v) & \longmapsto & \lambda\cdot v \end{array}$$

telles que :

- E soit un groupe commutatif pour la loi $+$; on note 0 son neutre (vecteur nul) ;
- les lois $+$ et $\cdot$ soient compatibles, c'est-à-dire :
  - $\forall v \in E, \quad 1\cdot v = v$;
  - $\forall\lambda,\mu \in \mathbb{K},\ \forall v \in E, \quad (\lambda\cdot\mu)\cdot v = \lambda\cdot(\mu\cdot v)$;
  - $\forall\lambda,\mu,\in \mathbb{K},\ \forall v \in E, \quad (\lambda+\mu)\cdot v = \lambda\cdot v+\mu\cdot v$;
  - $\forall\lambda \in \mathbb{K},\ \forall v,w \in E, \quad \lambda\cdot(v+w) = \lambda\cdot v+\lambda\cdot w$;

Exercice : Pour $\lambda \in \mathbb{K}$ et $v \in E$, $\lambda.v = 0 \Longleftrightarrow \lambda = 0$ ou $v = 0$. Par ailleurs, $\lambda\cdot(-v) = -\lambda\cdot v$.

Attention ! *L'inverse d'un vecteur v^{-1}, le produit d'un vecteur par un scalaire $v\cdot\lambda$, le produit de deux vecteurs $v \cdot w$, la somme d'un vecteur et d'un scalaire $\lambda + v$ n'ont en général aucun sens.*

Exemples d'espaces vectoriels

1. Si $\mathbb{K}$ est un corps et $\mathbb{L}$ est un corps qui contient $\mathbb{K}$, alors $\mathbb{L}$ est naturellement un espace vectoriel sur $\mathbb{K}$: la somme des vecteurs est celle de $\mathbb{L}$, le produit d'un scalaire par un vecteur est la restriction du produit de $\mathbb{L}$ à $\mathbb{K} \times \mathbb{L} \subset \mathbb{K} \times \mathbb{L}$. En particulier, $\mathbb{K}$ est un espace vectoriel sur $\mathbb{K}$.

2. Le singleton $\{0\}$ est un espace vectoriel sur tout corps $\mathbb{K}$.

3. Si E est un espace vectoriel sur $\mathbb{K}$ et X est un ensemble, alors l'ensemble E^X des fonctions de X dans E est naturellement un espace vectoriel sur $\mathbb{K}$. Opérations, pour $f, g \in E^X$ et $\lambda \in \mathbb{K}$:
$$\forall x \in X, \qquad (f+g)(x) = f(x) + g(x), \quad (\lambda \cdot f)(x) = \lambda \cdot f(x).$$

   Cas particuliers :

   (a) Si on prend $E = \mathbb{K}$, on voit que les fonctions à valeurs dans $\mathbb{K}$ forment un espace vectoriel sur $\mathbb{K}$.
   Par exemple, l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ (resp. $\mathbb{C}$) est naturellement un espace vectoriel sur $\mathbb{R}$ (resp. $\mathbb{C}$) ;

   (b) Si on prend $X = \{1, \ldots, n\}$ pour $n \in \mathbb{N}^*$, une fonction de X dans E est simplement une n-liste d'éléments de E : E^X s'identifie naturellement à E^n. En particulier, $\mathbb{K}^n$ rentre dans ce cadre.

   (c) Si on prend $E = \mathbb{K}$ et $X = \{1, \ldots, m\} \times \{1, \ldots, n\}$, où $m, n \in \mathbb{N}^*$, il est traditionnel de noter a_{ij} l'image de $(i, j) \in X$ par une fonction $A \in \mathbb{K}^X$. On obtient alors l'espace des matrices $m \times n$ à coefficients dans $\mathbb{K}$.

2° La sous-structure

a) Soit E un espace vectoriel sur un corps $\mathbb{K}$. On appelle sous-espace vectoriel de E toute partie F telle que
$$\forall u, u' \in F, \quad \forall \lambda, \lambda' \in \mathbb{K}, \qquad \lambda\, u + \lambda'\, u' \in F.$$

Exemples de sous-espaces vectoriels

1. Pour tout espace E, $\{0\}$ et E sont des sous-espaces. De plus, tout sous-espace F de E contient le vecteur nul, i.e. $\{0\} \subset F$.

2. L'ensemble des fonctions continues de $[0, 1]$ dans $\mathbb{R}$ est un sous-espace de l'espace des fonctions de $[0, 1]$ dans $\mathbb{R}$. L'ensemble des fonctions continues d'intégrale nulle aussi, mais pas l'ensemble des fonctions continuers d'intégrale 2 (qui est un sous-espace affine).

3. Pour E un espace vectoriel et X un ensemble, l'ensemble des fonctions à support fini est un sous-espace de l'espace E^X des fonctions de X dans E. On le note $E^{(X)}$. Ici, le support d'une fonction est l'ensemble des éléments de X dont l'image n'est pas nul. Noter que X est fini si et seulement si $E^X = E^{(X)}$ (pour peu que $E \neq \{0\}$).

   Par exemple, si $X = \mathbb{N}$, on sait que $\mathbb{R}^{(\mathbb{N})}$ s'identifie naturellement à l'espace des polynômes réels : une suite $(a_n)_{n\in\mathbb{N}}$ appartient à $\mathbb{R}^{(\mathbb{N})}$ si et seulement si elle est nulle à partir d'un certain rang, dans ce cas on peut l'interpréter comme la suite des coefficients d'un polynôme.

4. **Familles, familles presque nulles.** Soit E un espace vectoriel et I un ensemble. Il est assez courant de noter $(u_i)_{i\in I}$ un élément de E^I : cela signifie que u_i désigne l'image de $i \in I$. On parle alors de famille indexée par I, et on dit qu'une famille est presque nulle si elle appartient à $E^{(I)}$.

b) Opérations sur les sous-espaces

Lemme *Soit E_1, E_2 deux sous-espaces d'un espace vectoriel E sur $\mathbb{K}$. Alors $E_1 \cap E_2$ et*

$$E_1 + E_2 = \{u \in E,\ \exists u_1 \in E_1,\ \exists u_2 \in E_2,\ u = u_1 + u_2\}$$

sont des sous-espaces de E.

Attention ! *La réunion de deux sous-espace n'est pas, en général, un sous-espace. Plus précisément, c'est un sous-espace si et seulement si un des deux sous-espace contient l'autre.*

On dit que deux sous-espaces E_1 et E_2 de E sont en somme directe si leur intersection est réduite à $\{0\}$. On dit qu'ils sont supplémentaires si et seulement s'ils sont en somme directe et si leur somme est E.

Lemme *Deux sous-espaces E_1 et E_2 de E sont supplémentaires si et seulement si tout vecteur de E peut s'écrire de façon unique comme somme d'un vecteur de E_1 et d'un vecteur de E_2.*

Exemples : Les fonctions paires et impaires forment deux supplémentaires dans l'espace des fonctions réelles.

3° Les morphismes

a) Soit E, F deux espaces vectoriels sur $\mathbb{K}$. Une application $\varphi : E \to F$ est dite linéaire si

$$\forall u, u' \in E,\ \forall \lambda, \lambda' \in \mathbb{K}, \quad \varphi(\lambda\, u + \lambda'\, u') = \lambda\varphi(u) + \lambda'\varphi(u').$$

On note classiquement $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F. C'est un sous-espace de l'espace des fonctions de E dans F. Si $E = F$, on note $\mathcal{L}(E) = \mathcal{L}(E, E)$.

Lemme *La composée de deux applications linéaires est linéaire. De plus, si $\varphi, \varphi' \in \mathcal{L}(E, F)$, $\psi, \psi' \in \mathcal{L}(F, G)$ et $\lambda, \lambda' \in \mathbb{K}$, on a :*

$$\psi \circ (\lambda\varphi + \lambda'\varphi') = \lambda\psi \circ \varphi + \lambda'\psi \circ \varphi', \quad (\lambda\psi + \lambda'\psi') \circ \varphi = \lambda\psi \circ \varphi + \lambda'\psi' \circ \varphi.$$

b) Etant donné $\varphi \in \mathcal{L}(E, F)$, on appelle noyau et image de φ les sous-espaces

$$\operatorname{Ker} \varphi = \{u \in E,\ \varphi(u) = 0\} \subset E, \qquad \operatorname{Im} \varphi = \{v \in F,\ \exists u \in E,\ \varphi(u) = v\} \subset F.$$

Lemme *Soit $\varphi \in \mathcal{L}(E, F)$. Alors φ est injective si et seulement si* $\operatorname{Ker} \varphi = \{0\}$.

c) Construction d'applications linéaires

Lemme *Soit E, F des espaces vectoriels, E_1 et E_2 deux sous-espaces supplémentaires de E. Etant donné $\varphi_i \in \mathcal{L}(E_i, F)$ ($i = 1, 2$), il existe un unique $\varphi \in \mathcal{L}(E, F)$ tel que $\varphi|_{E_i} = \varphi_i$.*

Exemple : Si on définit $\varphi_1 : E_1 \to E$, $u \mapsto u$ et $\varphi_2 : E_2 \to E$, $u \mapsto 0$, l'application φ ainsi construite est la projection sur E_1 parallèlement à E_2. C'est un idempotent : elle est égale à son carré. De plus, tous les idempotents de $\mathcal{L}(E)$ sont de ce type.

4° Familles

a) Combinaisons linéaires

Soit E un espace vectoriel sur $\mathbb{K}$, $\mathcal{F} = (u_i)_{i \in I} \in E^I$ une famille quelconque d'éléments de E. On dit qu'un vecteur $u \in E$ est une combinaison linéaire de $\mathcal{F}$ s'il existe $(\lambda_i)_{i\in I} \in \mathbb{K}^{(I)}$ une famille presque nulle de scalaires telle que $u = \sum_{i\in I} \lambda_i\, u_i$.

Attention ! *La somme précédente n'aurait pas de sens si on n'imposait pas que seuls un nombre fini de λ_i soient non nuls.*

On est naturellement amené à considérer l'application

$$\begin{array}{rccc} c_{\mathcal{F}} : & \mathbb{K}^{(I)} & \longrightarrow & E \\ & (\lambda_i)_{i\in I} & \longmapsto & \sum_{i\in I} \lambda_i\, u_i. \end{array}$$

Il est facile de voir que cette application est linéaire. On voit que u est combinaison linéaire de $\mathcal{F}$ si et seulement si u est dans l'image de $c_{\mathcal{F}}$: ceci prouve donc que l'ensemble des combinaisons linéaires de $\mathcal{F}$ forme un sous-espace vectoriel. On montre par une récurrence lourde mais facile sur le nombre de coefficients non nuls de $(\lambda_i)_{i\in I}$ que si un sous-espace contient tous les u_i, alors il contient toutes leurs combinaisons linéaires. En d'autres termes :

Lemme *L'ensemble des combinaisons linéaires d'une famille de vecteurs est le plus petit sous-espace qui contient tous les vecteurs de la famille.*

On l'appelle espace vectoriel engendré par la famille, souvent noté $\mathrm{Vect}(\mathcal{F})$.

b) Familles libres, familles génératrices

Une famille $\mathcal{F} = (u_i)_{i\in I} \in E^I$ d'éléments de E est dite *génératrice* si tout vecteur est combinaison linéaire de $\mathcal{F}$, i.e. si $\mathrm{Vect}(F) = E$, i.e. si pour tout $u \in E$, il existe $(\lambda_i)_{i\in I}$ une famille presque nulle de scalaires telle que $u = \sum_{i\in I} \lambda_i\, u_i$.

Une famille $\mathcal{F} = (u_i)_{i\in I} \in E^I$ d'éléments de E est dite *libre* si la seule combinaison linéaire nulle de $\mathcal{F}$ est la combinaison linéaire triviale, i.e. si, étant donné $(\lambda_i)_{i\in I}$ une famille presque nulle de scalaires telle que $\sum_{i\in I} \lambda_i\, u_i = 0$, on a : $\lambda_i = 0$ pour tout $i \in I$.

Remarque *Une famille $\mathcal{F}$ est libre (resp. génératrice) si et seulement si l'application $c_{\mathcal{F}}$ est injective (resp. surjective).*

Chapitre 1
Bases, matrices

On fixe un corps $\mathbb{K}$.

1° Base d'un espace vectoriel

a) Base d'un espace vectoriel. Rappelons qu'une base $\mathcal{B} = (e_j)_{j\in I}$ d'un espace vectoriel E est une famille libre et génératrice. En d'autres termes, tout vecteur $v \in E$ peut s'écrire de façon unique comme combinaison linéaire des vecteurs de $\mathcal{B}$, c'est-à-dire sous la forme $v = \sum_{j\in I} x_j e_j$, où les x_j $(j \in J)$ sont des scalaires tous nuls sauf un nombre fini.
Commençons par quelques évidences :

Proposition *Toute famille libre maximale est génératrice. Toute famille génératrice minimale est libre.*

Ici, la maximalité ou la minimalité se rapporte à l'inclusion.
On montre la première assertion par l'absurde, en ajoutant un vecteur qui ne serait pas engendré pour obtenir une famille libre plus grande. Quant à la deuxième, on la montre en supprimant un vecteur qui aurait un coefficient non nul dans une combinaison linéaire nulle. (Détailler.)
L'existence de bases en découle : par exemple, on part d'une famille génératrice (l'espace entier s'il le faut !) et on en extrait une sous-famille génératrice minimale. En passant, on a tacitement mais lourdement utilisé l'axiome du choix. On peut être plus précis :

Théorème (de la base incomplète) *Soit $\mathcal{L}$ une famille libre et $\mathcal{G}$ une famille génératrice de E contenant $\mathcal{L}$. Alors il existe une base $\mathcal{B}$ de E telle que $\mathcal{L} \subset \mathcal{B} \subset \mathcal{G}$.*

Preuve. Soit $\mathbb{L}$ l'ensemble des familles de vecteurs $\mathcal{F}$ qui sont libres et vérifient $\mathcal{L} \subset \mathcal{F} \subset \mathcal{G}$. Cet ensemble n'est pas vide car il contient $\mathcal{L}$.
[Il est inductivement ordonné par l'inclusion : si $(\mathcal{F}_n)_{n\in\mathbb{N}}$ est une suite croissante dans $\mathbb{L}$, la réunion $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ est un élément de $\mathbb{L}$ qui majore les $\mathcal{F}_n$. Le lemme de Zorn s'applique donc.]
Soit $\mathcal{B}$ un élément maximal de $\mathbb{L}$. C'est donc une famille libre. S'il existait un élément $v \in \mathcal{G}$, $v \notin \mathrm{Vect}(\mathcal{B})$, alors $\mathcal{B} \cup \{v\}$ serait libre (vérifier), ce qui contredit la maximalité de $\mathcal{B}$. Ainsi, $\mathcal{B}$ est une base.□

Corollaire *Tout sous-espace possède un supplémentaire.*

Preuve. Fixons une base du sous-espace de départ. C'est une famille libre de l'espace entier, que l'on complète en une base. Le sous-espace engendré par les vecteurs que l'on a ajoutés est un supplémentaire de notre sous-espace. (Vérifier.) □

Théorème *Deux bases d'un espace vectoriel ont toujours le même cardinal.*

Preuve. On ne donne la preuve que dans le cas où une des bases est finie. Il suffit de montrer que si $(e_1, \ldots, e_m)$ est une base de E et si $(f_1, \ldots, f_n)$ est libre avec $n \geq m$, alors $n = m$.
On modifie la base petit à petit. Pour commencer, on sait que f_1 est une combinaison linéaire des $(e_i)_{i=1,\ldots,m}$, disons : $f_1 = a_1e_1 + \cdots + a_me_m$. L'un des coefficients a_i n'est pas nul, sans quoi on aurait $f_1 = 0$. Quitte à renuméroter les e_i, on peut supposer que $a_1 \neq 0$. Il est facile de vérifier que $\mathcal{B}_1 = (f_1, e_2, \ldots, e_m)$ est une nouvelle base.
Supposons savoir que $(f_1, \ldots, f_k, e_{k+1}, \ldots, e_m)$ est une base pour $1 \leq k \leq n$. Montrons qu'après avoir renuméroté les e_i $(i \geq k + 1)$ si nécessaire, $(f_1, \ldots, f_{k+1}, e_{k+2}, \ldots, e_m)$ est une base.
Par l'hypothèse de récurrence, f_{k+1} est combinaison linéaire des vecteurs de la base : $f_{k+1} = c_1\, f_1 + \cdots + c_k\, f_k + c_{k+1}\, f_{k+1} + \cdots + c_m\, e_m$. L'un des scalaires $c_{k+1}, \ldots, c_m$ n'est pas nul, sans

quoi on aurait une combinaison linéaire non triviale des f_ℓ ($\ell \leq k+1$). Quitte à renuméroter les e_i ($i \geq k+1$)), on peut supposer que $c_{k+1} \neq 0$. On termine facilement la récurrence.
A la dernière étape on obtient : $(f_1, \ldots, f_m)$ est une base, ce qui entraîne en particulier $m = n$ par la première proposition.□

Remarque : La preuve précédente montre en fait qu'une famille de $m+1$ vecteurs dans un espace engendré par m vecteurs est liée.

Dans la suite, on ne considère plus que des espaces vectoriels de dimension finie.

b) Coordonnées. Données : E un espace vectoriel, $\mathcal{B} = (e_j)_{j=1 \text{ à } n}$ une base. On peut alors définir un isomorphisme "coordonnées" $c = c_E : E \to \mathbb{K}^n$ ainsi. Un vecteur $v \in E$ s'écrit de façon unique sous la forme $v = \sum_{j=1}^n x_j e_j$, on pose : $c(v) = (x_j)_{j=1 \text{ à } n}$. La réciproque de c est clairement définie par : $c^{-1}((x_j)_{j=1 \text{ à } n}) = \sum_{j=1}^n x_j e_j$. On considèrera le plus souvent $c(v)$ comme un vecteur-colonne, dit "colonne des coordonnées de v dans $\mathcal{B}$".
Cette application c sert à transformer l'espace vectoriel "abstrait" E en un espace vectoriel "concret", $\mathbb{K}^n$.

2° Matrices d'une application linéaire

a) Définition. Données : E et F, deux espaces ; $\varphi : E \to F$ une application linéaire ; $\mathcal{B} = (e_j)_{j=1 \text{ à } n}$ une base de E, $\mathcal{C} = (f_i)_{i=1 \text{ à } m}$ une base de F. On considère la matrice A de φ dans les base $\mathcal{B}$ et $\mathcal{C}$, construite colonne par colonne selon la règle :

$$\boxed{\text{la } j\text{ème colonne de la matrice est la colonne des coordonnées dans } \mathcal{C} \text{ de } \varphi(e_j).}$$

On a donc :

$$\boxed{\varphi(e_j) = \sum_{i=1}^m a_{ij} f_i}.$$

Exemple : Etant donnée une matrice A de taille $m \times n$, on considère l'application linéaire $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$, $X \mapsto AX$. Soit $\mathcal{B}_{\text{can}}$ et $\mathcal{C}_{\text{can}}$ les bases canoniques de $\mathbb{K}^n$ et $\mathbb{K}^m$. Alors on a : $\text{Mat}_{\mathcal{B}_{\text{can}}, \mathcal{C}_{\text{can}}}(\varphi_A) = A$.

b) Calcul de l'image des vecteurs. Si $v = \sum_{j=1}^n x_j e_j$, on a par linéarité :

$$\varphi(v) = \sum_{j=1}^n x_j \varphi(e_j) = \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} f_i = \sum_{j=1}^n \sum_{i=1}^m x_j a_{ij} f_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right) f_i.$$

On constate donc que *la colonne des coordonnées de $\varphi(v)$ est le produit de A par la colonne des coordonnées de v.*

c) Deux isomorphismes fondamentaux. Données : E et F, $\mathcal{B} = (e_j)_{j=1 \text{ à } n}$, $\mathcal{C} = (f_i)_{i=1 \text{ à } m}$ comme ci-dessus. On note $\mathcal{M}_{m,n}(\mathbb{K})$ l'espace vectoriel des matrices de taille $m \times n$ et $\mathcal{L}(E, F)$ l'espace des applications linéaires de E dans F. On a déjà construit une application $\text{Mat}_{\mathcal{B},\mathcal{C}} : \mathcal{L}(E, F) \to \mathcal{M}_{m,n}(\mathbb{K})$. Inversement, étant donnée une matrice $A = (a_{ij})_{i,j} \in \mathcal{M}_{m,n}(\mathbb{K})$, on note $\Phi_{\mathcal{B},\mathcal{C}}(A)$ l'unique application linéaire telle que $[\Phi_{\mathcal{B},\mathcal{C}}(A)](e_j) = \sum_{i=1}^m a_{ij} e'_i$.

Proposition *Les applications* $\text{Mat}_{\mathcal{B},\mathcal{C}}$ *et* $\Phi_{\mathcal{B},\mathcal{C}}$ *sont des isomorphismes réciproques entre* $\mathcal{L}(E, F)$ *et* $\mathcal{M}_{m \times n}(\mathbb{K})$.

d) Interprétation. Etant donnée une matrice A de taille $m \times n$, on considère l'application linéaire $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$, $X \mapsto AX$. On peut alors considérer le diagramme :

$$\begin{tikzcd}
E \arrow[r, "\varphi"] \arrow[d, "c_E"'] & F \arrow[d, "c_F"'] & \text{étage ``abstrait''} \arrow[d, "\text{coordonnées}"] \\
\mathbb{K}^n \arrow[r, "\varphi_A"] & \mathbb{K}^m & \text{étage ``concret''}
\end{tikzcd}$$

La formule de b) s'interprète ainsi : $c_F \circ \varphi = \varphi_A \circ c_E$. L'application φ_A donne une "image" parfaitement fidèle de φ. Par exemple, on a la

Proposition (i) *c_E induit un isomorphisme entre* $\mathrm{Ker}\,\varphi$ *et* $\mathrm{Ker}\,\varphi_A$ *;*
(ii) *c_F induit un isomorphisme entre* $\mathrm{Im}\,\varphi$ *et* $\mathrm{Im}\,\varphi_A$*.*

Démonstration : (i) Soit $v \in E$. On a : $v \in \mathrm{Ker}\,\varphi$ SSI $\varphi(v) = 0$ SSI $c_F \circ \varphi(v) = 0$ SSI $\varphi_A \circ c_E(v) = 0$ SSI $c_E(v) \in \mathrm{Ker}\,\varphi_A$.
(ii) Soit $w \in F$. On a : $w = \varphi(v)$ SSI $c_F(w) = c_F \circ \varphi(v)$ SSI $c_F(w) = \varphi_A \circ c_E(v)$. Ainsi, $w \in \mathrm{Im}\,\varphi$ SSI $c_F(w) \in \mathrm{Im}\,\varphi_A$.

Morale : En termes vagues, l'application φ_A est une "version concrète" de l'application "abstraite" φ : on ne perd pas d'information en passant de φ à φ_A, mais les calculs sont généralement plus faciles à l'étage matriciel.
Inversement, étant donnée une matrice A, il est souvent utile de savoir qu'il existe un "objet abstrait intrinsèque" (=une application linéaire qui a pour matrice A) pour étudier A.

e) Composition des applications et produit de matrices. Soit E, F, G des espaces vectoriels munis de bases $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$, et $\varphi : E \to F$ et $\psi : F \to G$ des applications linéaires. Alors,

$$\boxed{\text{la matrice de la composée est le produit des matrices.}}$$

Plus précisément, on a : $\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(\psi \circ \varphi) = \mathrm{Mat}_{\mathcal{C},\mathcal{D}}(\psi)\,\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi)$.

3° Théorème du rang

On définit le rang d'une application linéaire φ comme la dimension de l'espace $\mathrm{Im}\,\varphi$. Le rang d'une matrice A est le rang de l'application φ_A : c'est donc la dimension de l'espace engendré par les colonnes de A. Le (ii) de la proposition ci-dessus montre que le rang d'une application linéaire est le rang de n'importe laquelle de ses matrices.

Attention ! *Pour le moment, rien ne dit que c'est la dimension de l'espace engendré par les lignes de A.*

Théorème *Soit $\varphi : E \to F$ linéaire.*
(i) *(Version abstraite.) La restriction de φ à tout supplémentaire du noyau de φ est une bijection sur l'image de φ. En particulier on a la formule :*

$$\dim \mathrm{Ker}\,\varphi + \mathrm{rg}\,\varphi = \dim E.$$

(ii) *(Version matricielle.) Il existe des bases $\mathcal{B}$ de E et $\mathcal{C}$ de F telles que*

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi) = \left(\begin{array}{c|c} I_r & 0_{r\times(n-r)} \\ \hline 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{array}\right),$$

où r est le rang de φ, I_r est la matrice identité $r \times r$ et les autres matrices sont nulles (en indice : leur format).

4° Changement de base

a) Etant données deux bases $\mathcal{B} = (e_i)_{i=1 \text{ à } n}$ et $\mathcal{B}' = (e'_i)_{i=1 \text{ à } n}$ d'un espace vectoriel E, on considère la matrice $P_{\mathcal{B},\mathcal{B}'}$ dite matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, constituée de la façon suivante :

Règle : Les colonnes de $P_{\mathcal{B},\mathcal{B}'}$ sont les coordonnées des vecteurs de $\mathcal{B}'$ exprimées dans $\mathcal{B}$.

Proposition *Si $v \in E$, on note $X \in \mathbb{K}^n$ sa colonne de coordonnées dans $\mathcal{B}$ et X' sa colonne de coordonnées dans $\mathcal{B}'$. Alors : $X = P_{\mathcal{B},\mathcal{B}'}X'$.*

Problème mnémotechnique : est-ce que c'est $X' = PX$ ou $X = PX'$? Réponse en regardant la première colonne de P : d'une part, c'est la colonne de coordonnées de e'_1, le premier vecteur de $\mathcal{B}'$, dans la base $\mathcal{B}$; d'autre part, c'est le produit de P par la colonne $(1, 0, \ldots, 0)$. On obtient ce que l'on veut : pour $v = e'_1$, on a $X' = (1, \ldots, 0)$ et X est la première colonne de $P_{\mathcal{B},\mathcal{B}'}$.

Remarque : Par définition, on a : $P_{\mathcal{B},\mathcal{B}'} = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{Id})$.

Corollaire $P_{\mathcal{B},\mathcal{B}'} = P^{-1}_{\mathcal{B}',\mathcal{B}}$.

b) Lien avec les applications linéaires.
Données : E, F deux espaces vectoriels ; $\varphi : E \to F$ linéaire ; $\mathcal{B}$, $\mathcal{B}'$ deux bases de E ; $\mathcal{C}$, $\mathcal{C}'$ deux bases de F.
On note alors : $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi)$, $A' = \mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(\varphi)$, $P = P_{\mathcal{B},\mathcal{B}'}$ et $Q = P_{\mathcal{C},\mathcal{C}'}$. On veut une relation entre ces matrices.
On sait que $\varphi = \mathrm{Id}_F \circ \varphi = \varphi \circ \mathrm{Id}_E$. Prenons les matrices de ces applications en considérant le diagramme :

$$\begin{tikzcd}
E, \mathcal{B}' \arrow[r, "\varphi"] \arrow[d, "\mathrm{Id}_E"] & F, \mathcal{C}' \arrow[d, "\mathrm{Id}_F"] & \text{étage “nouvelles bases”} \arrow[d, "\text{changement de base}"] \\
E, \mathcal{B} \arrow[r, "\varphi"] & F, \mathcal{C} & \text{étage “anciennes bases”}
\end{tikzcd}$$

On a donc : $\mathrm{Mat}_{\mathcal{B}',\mathcal{C}}(\varphi) = \mathrm{Mat}_{\mathcal{C}',\mathcal{C}}(\mathrm{Id}_F).\,\mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(\varphi) = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi).\,\mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{Id}_E)$, d'où $QA' = AP$, ou encore :

$$\boxed{A' = Q^{-1}\, A\, P.}$$

Etant données deux matrices A et A', s'il existe deux matrices inversibles P et Q telles que la relation ci-dessus soit satisfaite, on dit que A et A' sont équivalentes. Cela signifie donc que A et A' représentent la même application linéaire dans des bases convenables.

Cas particulier important : Si $E = F$, $\mathcal{B} = \mathcal{C}$ et $\mathcal{B}' = \mathcal{C}'$, on n'a qu'une seule matrice de changement de base : $P = Q = P_{\mathcal{B},\mathcal{B}'}$. Dans ce cas, on a : $A' = P^{-1}\, A\, P$. On dit que A et A' sont semblables. Cela signifie que A et A' représentent le même endomorphisme dans des bases convenables.

c) Théorème du rang (bis). En appliquant la version (ii) du théorème à φ_A, maintenant que l'on connait l'effet d'un changement de bases sur une matrice, on voit que pour toute matrice A, il existe deux matrices inversibles P et Q telles que $Q^{-1}AP$ soit la matrice du (ii).

d) Invariance du rang : Si $A' = Q^{-1}AP$, alors $\mathrm{rg}\, A = \mathrm{rg}\, A'$.
En effet, on interprète A et A' comme la matrice de la même application linéaire φ dans les bases dont les matrices de passage sont P et Q. Alors $\mathrm{rg}\, A$ et $\mathrm{rg}\, A'$ valent le rang de φ.

C'est le bon moment pour faire la fiche 4 : Algorithme de Gauss et avatars !

On note $E = \mathbb{K}^n$ et $F = \mathbb{K}^m$. On se donne une base $\mathcal{B}$ de $E = \mathbb{K}^n$ et $\mathcal{C}$ de $F = \mathbb{K}^m$.

applications linéaires	matrices	systèmes linéaires
$\varphi : \mathbb{R}^n \to \mathbb{R}^m$ linéaire	$A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi)$	systèmes $AX = B$ (B parcourt $\mathbb{K}^m$)
noyau de l'application linéaire	noyau de la matrice	solutions du système $AX = 0$
image de l'application linéaire	espace engendré par les colonnes	ensemble des $B \in \mathbb{R}^m$ tel que $AX = B$ a une solution
rang de l'application = dimension de l'image	rang de la matrice = dimension de l'espace des colonnes	rang du système = nombre d'inconnues principales
dimension du noyau	dimension du noyau	$n - \mathrm{rg} =$ nombre d'inconnues secondaires
modifier $\mathcal{B}$	multiplier A à droite par une matrice inversible opérations élémentaires sur les colonnes	changer les inconnues
modifier $\mathcal{C}$	multiplier A à gauche par une matrice inversible opérations élémentaires sur les lignes	opérations élémentaires sur les équations

Chapitre 2
Dualité, transposition

On fixe un corps $\mathbb{K}$. On ne considère que des espaces vectoriels de dimension finie.

1° Dualité

a) Base duale. Le dual d'un espace vectoriel E est simplement : $E^* = \mathcal{L}(E, \mathbb{K})$. Etant donné une base $\mathcal{B} = (e_i)_{i=1,\ldots,n}$ de E, on définit une famille $\mathcal{B}^* = (e_i^*)_{i=1,\ldots,n}$ de E^* par l'image de $\mathcal{B}$:

$$e_i^*(e_j) = \delta_{ij} \qquad (i, j = 1, \ldots, n).$$

Un vecteur $v \in E$ s'écrit de façon unique $v = \sum_{j=1}^n x_j e_j$, pour $(x_j)_{j=1,\ldots,n} \in \mathbb{K}^n$. On a donc :

$$e_i^*(v) = e_i^*\left(\sum_{j=1}^n x_j e_j\right) = x_i, \qquad \text{d'où} \quad \boxed{v = \sum_{i=1}^n e_i^*(v)\, e_i.}$$

En mots, $e_i^*(v)$ est la $i$ème coordonnée de v dans la base $\mathcal{B}$.
Remarque : sans donnée supplémentaire, un vecteur $x \in E$ *ne suffit pas* à définir une application linéaire $x^* \in E^*$. Cependant, étant donnée une base $\mathcal{B}$ de E, on définit sans ambiguïté une application linéaire $\tau : E \to E^*$ (qui dépend de $\mathcal{B}$) par : $\tau(e_i) = e_i^*$, qui est clairement un isomorphisme.

Proposition *$\mathcal{B}^*$ est une base de E^*.*

Définition : On dit que $\mathcal{B}^*$ est la base duale de $\mathcal{B}$.
Remarque : Si E est de dimension infinie, $\mathcal{B}^*$ est encore définie, libre mais pas génératrice.

b) Base biduale. Soit E un espace vectoriel de dimension finie.
D'une part, on peut définir une application linéaire $\iota : E \to E^{**}$ par : $[\iota(v)](\ell) = \ell(v)$ pour $\ell \in E^*$. C'est un isomorphisme (car on suppose E de dimension finie, sinon ce serait seulement une injection) et il ne dépend d'aucun choix : désormais, on *identifiera* E et E^{**}.
D'autre part, on peut considérer la base duale de $\mathcal{B}^*$. C'est la base $\mathcal{B}^{**} = (e_j^{**})_{j \in I}$ définie par : $e_j^{**}(e_i^*) = \delta_{ij}$. On a donc : $e_j^{**}(e_i^*) = [\iota(e_j)](e_i^*)$, soit $\mathcal{B}^{**} = \iota(\mathcal{B})$.

Morale : Comme on décide d'identifier E et E^{**}, la base duale de $\mathcal{B}^*$ est donc $\mathcal{B}$.

c) $\mathbb{K}^n$, c'est des lignes ou des colonnes ? Considérons l'espace vectoriel Col_n des matrices $n \times 1$ et l'espace Lign_n des matrices $1 \times n$. Etant donnée une ligne $L \in \mathrm{Lign}_n$, considérons la forme linéaire $\mathrm{Col}_n \to \mathbb{K}$, $C \mapsto LC$. Puisque LC est une matrice 1×1, c'est un scalaire. De cette façon, on identifie Lign_n à Col_n^*, le dual de Col_n.
Dans ce contexte, l'application $\tau : \mathrm{Col}_n \to \mathrm{Col}_n^* = \mathrm{Lign}_n$ associée à la base naturelle de Col_n, c'est la transposition (qui transforme les colonnes en lignes et inversement).
Remarque : On a aussi une base naturelle dans Lign_n. La colonne des coordonnées de $L \in \mathrm{Lign}_n$ dans cette base est simplement : ${}^tL$, la transposée de L.

Cela dit, cette distinction est un peu artificielle, vu qu'étant donné un vecteur $X \in \mathbb{K}^n$, on peut en faire un ligne ou une colonne. Cela veut dire qu'on peut identifier $\mathbb{K}^n$ et son dual de façon standard. Si $X' \in \mathbb{K}^n$, on peut le considérer comme la forme linéaire $\mathbb{K}^n \to \mathbb{K}$, $X \mapsto {}^tX'\, X$.
Concrètement, quand on veut parler de produit par une matrice, il est commode d'identifier un vecteur de $\mathbb{K}^n$ à la colonne de ses coordonnées dans la base canonique.
Toutes ces considérations sont dues au fait que $\mathbb{K}^n$ possède une base canonique, ce qui n'est pas le cas d'un espace vectoriel quelconque.

Application : Si $v \in E$ a pour coordonnées $X \in \mathbb{K}^n$ dans une base $\mathcal{B}$ et si $\ell \in E^*$ a pour coordonnées $X' \in \mathbb{K}^n$ dans la base duale $\mathcal{B}^*$, alors : $\ell(v) = {}^tX'\, X$ (c'est une matrice 1×1, donc un scalaire).

2° Transposition

a) Définitions, etc. On se donne une application linéaire $\varphi : E \to F$ entre espaces vectoriels. On définit alors une application ${}^t\varphi : F^* \to E^*$ par : ${}^t\varphi(\ell) = \ell \circ \varphi$ pour $\ell \in F^*$ (c'est-à-dire : ${}^t\varphi(\ell)(v) = \ell(\varphi(v))$ pour $v \in E$).

Proposition (i) *L'application $T : \mathcal{L}(E, F) \to \mathcal{L}(F^*, E^*)$, $\varphi \mapsto {}^t\varphi$ est linéaire.*
(ii) *Si E et F sont de dimension finie, on identifie $\mathcal{L}(E^{**}, F^{**})$ à $\mathcal{L}(E, F)$ et alors : ${}^t({}^t\varphi) = \varphi$. En particulier, dans ce cas, T est un isomorphisme.*
(iii) *Si $\varphi : E \to F$ et $\psi : F \to G$, on a : ${}^t(\psi \circ \varphi) = {}^t\varphi \circ {}^t\psi$ (dans $\mathcal{L}(G^*, E^*)$).*

b) Exemple. Soit A une matrice $m \times n$, considérons $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$, $X \mapsto AX$. (Ici et dans la suite, on identifie $\mathbb{K}^n$ et Col_n ; de plus on choisit de travailler avec la base canonique de $\mathbb{K}^n$, ce qui permet d'identifier $\mathbb{K}^n$ à son dual ; de même avec $\mathbb{K}^m$.)
Prenons $\ell = Y$ dans le dual de $\mathbb{K}^m$ et $w = Y' \in \mathbb{K}^m$. Alors, $\ell(w)$ s'écrit ${}^tY\,Y'$. En particulier, si w est de la forme $\varphi_A(v) = AX$ pour $v = X \in \mathbb{K}^n$, on a : $\ell(\varphi_A(v)) = {}^tYAX$. Or, par définition : $\ell(\varphi(v)) = [{}^t\varphi_A(\ell)](v)$ et ${}^tYAX = {}^t({}^tAY)X$. Ceci entraîne que : $[{}^t\varphi_A(\ell)] = {}^tAY$ (vérifier !). En d'autres termes ($\ell = Y$), on a, modulo les identifications évoquées ci-dessus :

$$ {}^t\varphi_A = \varphi_{{}^tA}. $$

3° Matrices

Soit $\mathcal{B} = (e_j)_{j=1 \text{ à } n}$ une base de E, $\mathcal{C} = (f_i)_{i=1 \text{ à } m}$ une base de F et $\varphi : E \to F$ linéaire.

a) Matrice d'une application et dualité. Considérons la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi)$. On la notera $A = (a_{ij})_{i,j}$ (pour $i = 1$ à m, $j = 1$ à n). On sait que les colonnes de A sont les coordonnées (dans $\mathcal{C}$) des images des vecteurs de $\mathcal{B}$:

$$ \varphi(e_j) = \sum_{i=1}^{m} a_{ij} e_j \quad (j = 1 \text{ à } n) \quad \text{d'où} \quad \boxed{a_{ij} = f_i^*\Big(\varphi(e_j)\Big).} $$

En mots : a_{ij}, le coefficient d'indices (i, j) de A, est le $i$ème coefficient de $\varphi(e_j)$.

Application : Si $v \in E$ a pour colonne de coordonnées $X \in \mathbb{K}^n$, alors $\varphi(v)$ a pour colonne de coordonnées $AX \in \mathbb{K}^m$. Grâce à la formule de 1°a), on voit que si par ailleurs $\ell \in F^*$ a pour coordonnées $Y \in \mathbb{K}^m$, on a : $\ell[\varphi(v)] = {}^tYAX$.

b) Matrice de la transposée.

Proposition $\mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t\varphi) = {}^t\,\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi)$.

Démonstration : Tout d'abord, le format de $A' = \mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t\varphi)$ est bien $n \times m$. On doit calculer ${}^t\varphi(f_i^*)$, car les coefficients a'_{ij} de A sont définis par :

$$ (\S) \qquad {}^t\varphi(f_i^*) = \sum_{i=1}^{n} a'_{ji} e_i^*. $$

<u>Première version</u> (directe) : L'application linéaire ${}^t\varphi(f_i^*) : E \to \mathbb{K}$ est déterminée par l'image de la base $\mathcal{B}$. On calcule donc :

$$ \begin{aligned} [{}^t\varphi(f_i^*)](e_j) &= f_i^*\Big(\varphi(e_j)\Big) && \text{par définition de } {}^t\varphi \\ &= f_i^*\left(\sum_{k=1}^{m} a_{kj} f_k\right) && \text{par définition de } (a_{ij})_{i,j} \\ &= a_{ij} && \text{par définition de la base duale } \mathcal{C}^*. \end{aligned} $$

Ceci signifie que ${}^t\varphi(f_i^*) = \sum_{j=1}^n a_{ij} e_j^*$, et donc la $i$ème colonne de A' est $(a_{ij})_{j=1 \text{ à } m}$, c'est-à-dire la $i$ème ligne de A.
Deuxième version (avec la base biduale) : On a donc :

$$\begin{array}{rcll} a'_{ji} & = & e_j^{**}({}^t\varphi(f_i^*)) & \text{par (§)} \\ & = & [{}^t\varphi(f_i^*)](e_j) & \text{par définition de } e_j^{**} \\ & = & f_i^*\Big(\varphi(e_j)\Big) & \text{par définition de } {}^t\varphi \\ & = & a_{ij} & \text{par } \mathbf{a)}. \end{array}$$

4° Application au rang

Rappelons que le rang d'une application linéaire est la dimension de son image ; le rang d'une matrice est la dimension de l'espace vectoriel engendré par ses colonnes dans Col_n ; le rang d'une application linéaire est le rang de n'importe quelle matrice qui la représente.

Proposition (i) *Soit* $\varphi : E \to F$ *linéaire. Alors* $\mathrm{rg}\,\varphi = \mathrm{rg}\,{}^t\varphi$.
(ii) *Soit* A *une matrice* $m \times n$. *Alors* $\mathrm{rg}\,A = \mathrm{rg}\,{}^tA$.

Démonstration : (i) 1ᵉ version : On sait que le rang de φ est le rang de n'importe quelle matrice qui représente φ. Avec le théorème du rang, on choisit des bases $\mathcal{B}$ et $\mathcal{C}$ de manière à ce que

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\varphi) = \left(\begin{array}{c|c} I_r & 0_{r\times(n-r)} \\ \hline 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{array}\right) = B.$$

Alors, le rang de B est $r = \mathrm{rg}\,\varphi$. On constate que $\mathrm{rg}\,B = \mathrm{rg}\,{}^tB$. Or, ${}^tB$ est la matrice de ${}^t\varphi$ (dans les bases $\mathcal{C}^*$ et $\mathcal{B}^*$). On a donc $\mathrm{rg}\,{}^t\varphi = \mathrm{rg}\,{}^tB = \mathrm{rg}\,B = \mathrm{rg}\,\varphi$.
2ᵉ version : Rappelons que l'orthogonal d'un sous-espace $G \subset F$ est $G^\perp = \{\ell \in F^*,\ \forall w \in G, \ell(w) = 0\}$: c'est un sous-espace de F^* de dimension $\dim G^\perp = \dim F - \dim G$. Or, on a : $(\mathrm{Im}\,\varphi)^\perp = \mathrm{Ker}\,{}^t\varphi$. En effet, on a :

$$\ell \in (\mathrm{Im}\,\varphi)^\perp \Leftrightarrow \forall v \in E,\ \ell(\varphi(v)) = 0 \Leftrightarrow \forall v \in E,\ [{}^t\varphi(\ell)](v) = 0 \Leftrightarrow {}^t\varphi(\ell) = 0 \Leftrightarrow v \in \mathrm{Ker}\,{}^t\varphi.$$

On en déduit : $\mathrm{rg}\,{}^t\varphi = \dim F - \dim \mathrm{Ker}\,{}^t\varphi = \dim F - \dim(\mathrm{Im}\,\varphi)^\perp = \dim \mathrm{Im}\,\varphi = \mathrm{rg}\,\varphi$.
(ii) 1ᵉ version : Considérons A comme la matrice de l'application linéaire φ_A dans les bases canoniques. On a : $\mathrm{rg}\,A = \mathrm{rg}\,\varphi_A = \mathrm{rg}\,{}^t\varphi_A = \mathrm{rg}\,\varphi_{{}^tA} = \mathrm{rg}\,{}^tA$.
2ᵉ version : D'après une des versions du théorème du rang, il existe P et Q des matrices inversibles telles que $Q^{-1}AP$ est la matrice B ci-dessus. On voit immédiatement que le rang de B est le rang de ${}^tB$. Or, multiplier une matrice par une matrice inversible ne change pas son rang. On a donc : $\mathrm{rg}\,A = \mathrm{rg}(QBP^{-1}) = \mathrm{rg}\,B = \mathrm{rg}\,{}^tB = \mathrm{rg}({}^tP\,{}^tA\,{}^tQ^{-1}) = \mathrm{rg}\,A$.

Corollaire *Le rang d'une matrice* $m \times n$ *est la dimension de l'espace vectoriel engendré par ses colonnes dans* Col_n *ou de ses lignes dans* Lign_m.

Il est encore temps de faire la fiche 4 : Algorithme de Gauss et avatars !

Chapitre 3
Déterminants

I La théorie

On fixe un espace vectoriel E sur un corps $\mathbb{K}$.

1° Formes multilinéaires alternées

a) Définitions : Soit n un entier et $f : E^n \to \mathbb{K}$. On dit que f est n-linéaire ou multilinéaire si elle est linéaire par rapport à chaque variable : pour tout $i = 1, \dots, n$ et toute famille $(v_j)_{j \neq i}$, on demande que l'application $f_i : E \to \mathbb{K}$, $v_i \mapsto f(v_1, \dots, v_n)$ soit linéaire.
On dit que f est alternée si elle s'annule lorsque deux variables prennent la même valeur : si $v_j = v_j$ avec $i \neq j$, alors $f(v_1, \dots, v_n) = 0$.

Lemme *Si f est alternée, on change le signe en permutant deux variables :*

$$\forall (v_i)_{i=1,\dots,n} \in E^n, \ \forall i \neq j, \quad f(\dots, v_i, \dots, v_j, \dots) = -f(\dots, v_j, \dots, v_i, \dots).$$

b) Calcul fondamental. On se donne deux familles $\mathbf{v} = (v_i)_{i=1,\dots,n}$ et $\mathbf{w} = (w_i)_{i=1,\dots,n}$ et une matrice $A = (a_{ij})_{i,j=1,\dots,n} \in \mathcal{M}_n(\mathbb{K})$ telle que

$$(*) \qquad \forall i = 1, \dots, n, \quad w_i = a_{i1}v_1 + \dots + a_{in}v_n.$$

On veut calculer $f(\mathbf{w})$ en fonction de $f(\mathbf{v})$. Pour cela, on développe en utilisant la multilinéarité de f. La $i$ème variable w_i est une somme de n termes $a_{i,j}v_j$; dans chacune on choisit un terme $a_{i,\sigma(i)}v_{\sigma(i)}$. On obtient une somme indexée par tous les choix possibles, i.e. toutes les applications $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$. En formule :

$$f(\mathbf{w}) = \sum_{\sigma : \{1,\dots,n\} \to \{1,\dots,n\}} a_{1,\sigma(1)} \dots a_{n,\sigma(n)} \, f(v_{\sigma(1)}, \dots, v_{\sigma(n)}).$$

Or, si σ n'est pas bijective, elle n'est pas injective : il existe $i \neq j$ tels que $\sigma(i) = \sigma(j)$. Le terme correspondant dans la somme ci-dessus est donc nul, puisque f est alternée.
Notons $\mathfrak{S}_n$ l'ensemble des bijections de $\{1, \dots, n\}$ dans $\{1, \dots, n\}$. On a donc :

$$f(\mathbf{w}) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \, a_{1,\sigma(1)} \dots a_{n,\sigma(n)} \, f(\mathbf{v}),$$

où, pour $\sigma \in \mathfrak{S}_n$, on note $\varepsilon(\sigma)$ sa signature.

c) Résultat technique. Supposons que E soit de dimension n et soit $\mathbf{v} = (v_i)_{i=1,\dots,n}$ une base de E. Toute famille de n vecteurs $\mathbf{w} = (w_i)_{i=1,\dots,n}$ détermine une unique matrice A de taille $n \times n$ par les équations $(*)$ ci-dessus. On pose alors :

$$\det_{\mathbf{v}}(\mathbf{w}) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \, a_{1,\sigma(1)} \dots a_{n,\sigma(n)}.$$

Lemme *Pour toute base $\mathbf{v} = (v_i)_{i=1,\dots,n}$ de E, l'application $\det_{\mathbf{v}}$ est n-linéaire et alternée. De plus, on a : $\det_{\mathbf{v}}(\mathbf{v}) = 1$.*

Démonstration. Admis, beaucoup d'indices à manipuler.□

d) Résultat principal.

Théorème *Supposons E de dimension n. Alors l'ensemble des formes n-linéaires alternées est un espace vectoriel de dimension 1.*

Démonstration. Notons F l'espace vectoriel des formes n-linéaires alternées et fixons une base $\mathbf{v}$ de E. L'application $\mathrm{eval}_{\mathbf{v}} : F \to \mathbb{K}$, $f \mapsto f(\mathbf{v})$ est linéaire. Le calcul fondamental montre qu'elle est injective. L'existence de la forme $\det_{\mathbf{v}}$ dans le résultat technique entraîne la surjectivité.□

e) Critère d'indépendance linéaire.

Lemme *Soit $\mathbf{v} \in E^n$ une base de E et $\mathbf{w} \in E^n$ une famille quelconque. Alors $\mathbf{w}$ est une base de E si et seulement si $\det_{\mathbf{v}}(\mathbf{w}) \neq 0$.*

Démonstration. Supposons d'abord que $\mathbf{w}$ soit liée. Alors un des w_i est combinaison linéaire des autres, et la n-linéarité de $\det_{\mathbf{v}}$ entraîne l'annulation de $\det_{\mathbf{v}}(\mathbf{w})$. (Vérifiez !)
Supposons à présent que $\mathbf{w}$ soit libre. D'après le lemme, on dispose de deux formes linéaires alternées non nulles $\det_{\mathbf{v}}$ et $\det_{\mathbf{w}}$. D'après le théorème, elles sont proportionnelles (et le coefficient est non nul !) : il existe $\alpha \in \mathbb{K}^*$ tel que $\det_{\mathbf{v}} = \alpha \det_{\mathbf{w}}$. En évaluant en $\mathbf{w}$, on voit que $\det_{\mathbf{v}}(\mathbf{w}) = \alpha \neq 0$.□

2° Déterminants des matrices

a) Définition. Pour $A = (a_{ij})_{i,j=1,\dots,n}$ une matrice carrée $n \times n$, on pose :

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)\, a_{1,\sigma(1)} \dots a_{n,\sigma(n)}.$$

Remarque. Si on considère le déterminant comme une fonction des n colonnes de A, chaque colonne étant un vecteur dans $\mathbb{K}^n$, on constate que det est l'unique forme n-linéaire alternée sur les colonnes d'une matrice qui prend la valeur 1 en Id.

b) Principales propriétés.

Proposition *Soit A et B deux matrices.*
(i) *A est inversible si et seulement si $\det(A) \neq 0$.*
(ii) $\det(A) = \det({}^tA)$.
(iii) $\det(AB) = \det A \ \det B$.

Attention ! *Le déterminant de la somme de deux matrices est différent de la somme des déterminants (en général) !*

Démonstration. (i) Ce n'est qu'une reformulation du lemme précédent pour les matrices, en se rappelant que la matrice A est inversible si et seulement si l'image de l'application $X \mapsto AX$ est $\mathbb{K}^n$, si et seulement si les colonnes de A forment une base de $\mathbb{K}^n$.
(ii) Par calcul direct : repose sur le fait que pour $\sigma \in \mathfrak{S}_n$, on a : $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$.
(iii) Remarquons d'abord que les colonnes de AB sont des combinaisons linéaires des colonnes de A. Alors, si $\det(A) = 0$, on en déduit que les colonnes de A sont liées, puis que les colonnes de AB sont liées, donc que $\det(AB) = 0$. En raisonnant de même avec les lignes (ce qui est licite depuis la démonstration de (ii)), on obtient que si $\det(B) = 0$, alors $\det(AB) = 0$.
Maintenant, observons que l'application $A \mapsto \det(AB)$ est linéaire par rapport aux colonnes de A. Donc il existe un scalaire $f(B)$, qui ne dépend que de B, tel que pour tout A on ait : $\det(AB) = f(B) \det(A)$. De même on montre l'existence de $g(A) \in \mathbb{K}$ tel que pour tout B on ait : $\det(AB) = g(A) \det(B)$.

Mais alors, dès que $\det(A)$ et $\det(B)$ ne sont pas nuls, on a $g(A)/\det(A) = f(B)/\det(B)$: ce scalaire ne dépend donc ni de A ni de B. On vient de montrer l'existence de $c \in \mathbb{K}$ tel que pour tout (A, B) avec $\det(A)\det(B) \neq 0$, on ait :

$$(\S) \qquad \det(AB) = c \det(A) \det(B).$$

En prenant $A = B = \mathrm{Id}$, on obtient $c = 1$, et la première partie de la démonstration montre que l'égalité (§) s'étend pour A et B quelconques.□

3° Calcul des déterminants

Les principales méthodes de calculs sont :

a) Manipulations sur les lignes et les colonnes : pour $i \neq j$ et $\alpha \in \mathbb{K}$, les opérations $C_i \leftarrow C_i + \alpha C_j$ et $L_i \leftarrow L_i + \alpha L_j$ ne changent pas le déterminant ; les transpositions $C_i \leftrightarrow C_j$ et $L_i \leftrightarrow L_j$ ont pour effet de multiplier le déterminant par (-1) ; les substitutions $C_i \leftarrow \alpha C_i$ et $L_i \leftarrow \alpha L_i$ ont pour effet de multiplier le déterminant par α (si $\alpha \neq 0$). (Pourquoi ? par multiplicativité du déterminant, bien sûr !)

b) Développement selon une rangée. Notons A_{ij} la matrice $(n-1) \times (n-1)$ obtenue à partir de A en supprimant la $i$ème ligne et la $j$ème colonne et notons $\mathrm{com}(A) = b_{ij}$ la matrice de coefficient général $b_{ij} = (-1)^{i+j} \det(A_{ij})$. On prouve :

Proposition *On a, pour $i, j = 1, \ldots, n$:*

$$\begin{aligned} \det A &= \sum_{k=1}^{n} a_{ik} b_{ik} \quad \textit{(développement selon la } i\textit{ème ligne)} \\ &= \sum_{k=1}^{n} a_{kj} b_{kj} \quad \textit{(développement selon la } j\textit{ème colonne)} \end{aligned}$$

Démonstration. Par multilinéarité, il suffit de prouver la formule lorsqu'un seul des coefficients de la rangée correspondante est non nul. Quitte à permuter des rangée, on peut supposer qu'on veut développer par rapport à la dernière rangée et que c'est le dernier coefficient (d'indice (n, n), donc) qui n'est pas nul. Alors, l'égalité résulte de la formule du déterminant d'une matrice.□

Les formules de développement se résument par une double égalité :

$$A\,{}^t\mathrm{com}(A) = \det(A)\,\mathrm{Id} = {}^t\mathrm{com}(A)\,A.$$

En effet, regardons le coefficient d'indices (i, j) des matrices ci-dessus. Pour $i = j$, la première égalité traduit le développement du déterminant de A selon la $i$ème ligne, la seconde selon la $i$ème colonne.

Pour $i \neq j$, le coefficient est nul car on calcule le déterminant de la matrice obtenue en remplaçant une rangée par une autre, et en développant selon cette rangée.

Il faut savoir calculer à vue certains déterminants :

c) Si la matrice est triangulaire, le déterminant est le produit des coefficients diagonaux.

d) Plus généralement, si la matrice est triangulaire par blocs, le déterminant est le produit des déterminants des blocs diagonaux.

Idée : La stratégie consiste donc en général à faire apparaître habilement des zéros sur une rangée (l'idéal est que tous les coefficients soient nuls sauf 1 ou 2), puis à développer le long de cette rangée.

II Des applications

1° Formules de Cramer

Soit $A \in \mathcal{M}_n(\mathbb{K})$, $B \in \mathbb{K}^n$. On s'intéresse au système $AX = B$. Les formules de Cramer donnent une réponse (rarement intéressante en pratique) lorsque A est inversible, ce que l'on suppose désormais.

Proposition *Soit $A \in \mathcal{M}_n(\mathbb{K})$ inversible, $B \in \mathbb{K}^n$. L'unique $X = (x_i)_{i=1\ldots,n} \in \mathbb{K}^n$ tel que $AX = B$ est donné par :*

$$x_i = \frac{\det A_i}{\det A},$$

où A_i est la matrice obtenue en remplaçant la ième colonne de A par B.

Exemple : si $n = 2$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $B = \begin{pmatrix} u \\ v \end{pmatrix}$, alors $X = \dfrac{1}{ad - bc}\begin{pmatrix} ud - bv \\ av - uc \end{pmatrix}$.

Démonstration. Aimable variante de la formule $A^{-1} = \frac{1}{\det A}\,{}^t\mathrm{com}(A)$. Exercice.

2° Théorème de Rouché-Fontené

Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$, $B \in \mathbb{K}^m$, $X \in \mathbb{K}^n$. On veut savoir quand le système $AX = B$ (m équations, n inconnues) possède une solution. Le théorème de Rouché-Fontené donne un critère en termes de déterminants.

Pour préparer un peu le terrain, on note $r = \mathrm{rg}(A)$ le rang du système, et on suppose que la sous-matrice A_0 de taille $r \times r$ en haut à gauche de A est inversible. (Quitte à permuter des variables et des équations, i.e. des lignes et des colonnes de A, on peut toujours se ramener à ce cas.) Le sous-système formé des r premières équations admet toujours une solution (cf. 1° !). On a donc $m - r$ “équations en trop”.

Pour $i = r + 1, \ldots, m$, on forme une matrice M_i de la façon suivante : on met A_0 en haut à gauche ; on met les r premières coordonnées de B en haut à droite, le début de la $i$ème ligne de A en bas à gauche, et b_i en bas à droite :

$$M_i = \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1r} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{r1} & \ldots & a_{rr} & b_r \\ \hline a_{i1} & \ldots & a_{ir} & b_i \end{array}\right) \qquad (i = r + 1, \ldots, m).$$

Proposition *(Notations et hypothèses ci-dessus.) Le système $AX = B$ possède une solution si et seulement si $\det M_i = 0$, pour $i = r + 1, \ldots, m$.*

Démonstration. D'abord, on se convainc que les $n - r$ dernières inconnues ne servent à rien. On les oublie, ce qui revient à supposer $r = n$. (Exercice : donner un sens à ces deux phrases !) Puis on applique l'algorithme de Gauss aux r premières équations, en s'autorisant à permuter les colonnes (ce qui revient à renuméroter les inconnues) et toutes les opérations usuelles sur les lignes. Puisque le bloc correspondant de A est inversible, on peut le mener à terme. Ceci revient à supposer que $A_0 = \mathrm{Id}$.

On se retrouve avec un système et une matrice de la forme

$$\left\{\begin{array}{llll} x_1 & & & = b'_1 \\ & x_2 & & = b'_2 \\ & \ldots & & \\ & & x_r & = b'_r \\ a_{i1}x_1 & + \ldots & + a_{ir}x_r & = b_i \end{array}\right. \quad \text{et} \quad M'_i = \left(\begin{array}{ccc|c} 1 & \ldots & 0 & b'_1 \\ \vdots & & \vdots & \vdots \\ 0 & \ldots & 0 & b'_r \\ \hline a_{i1} & \ldots & a_{ir} & b_i \end{array}\right) \quad (i = r + 1, \ldots, m).$$

Vu les manipulations qu'on a faites, on a : $\det M_i = 0$ si et seulement si $\det M'_i = 0$. Il n'y a plus qu'à vérifier que l'on a $\det M'_i = \sum_{j=1}^{r} a_{ij} b'_j - b'_i$, et que c'est bien la relation de compatibilité. C'est trivial.□

3° Orientations d'un espace vectoriel réel

Soit E un espace vectoriel réel de dimension finie non nulle.

Définition *On dit qu'une base $\mathcal{B}$ définit la même orientation de E que la base $\mathcal{B}'$ si le déterminant de la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est strictement positif.*

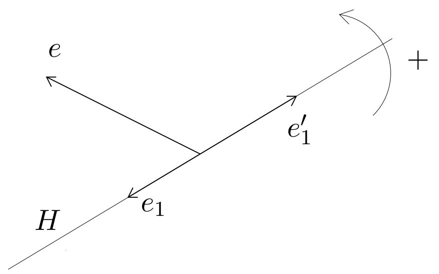
Lemme (i) *La relation "définir la même orientation de E que" est une relation d'équivalence sur l'ensemble des bases de E.*
(ii) *Cette relation possède exactement deux classes d'équivalence.*

Définition *Une orientation de E est une des deux classes d'équivalence de la relation "définir la même orientation". On dit que E est orienté si on a choisi une orientation. Les bases directes sont alors les bases qui appartiennent à cette orientation.*

Exemple : Si $E = \mathbb{R}^n$, on choisit généralement l'orientation de la base canonique/standard.
Démonstration. On va noter $P_{\mathcal{B},\mathcal{B}'} = M_{\mathcal{B}',\mathcal{B}}(\mathrm{Id})$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.
(i) La matrice de passage de $\mathcal{B}$ à $\mathcal{B}$ est l'identité, donc son déterminant est positif et la relation est réflexive.
La matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$ est l'inverse de la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, donc toutes deux ont un déterminant de même signe. La relation est donc symétrique.
On a l'égalité : $P_{\mathcal{B},\mathcal{B}''} = P_{\mathcal{B}',\mathcal{B}''} P_{\mathcal{B},\mathcal{B}'}$. Par conséquent, d'après la multiplicativité du déterminant, si $\mathcal{B}$ définit la même orientation que $\mathcal{B}'$ et $\mathcal{B}'$ définit la même orientation que $\mathcal{B}''$, alors le déterminant de $P_{\mathcal{B},\mathcal{B}''}$ est positif. La relation est donc transitive.
(ii) Si $\mathcal{B} = (e_1, \ldots, e_n)$ est une base de E, alors $\mathcal{B}$ et $\mathcal{B}^{\mathrm{op}} = (e_1, \ldots, e_{n-1}, -e_n)$ ne définissent pas la même orientation, donc il y a au moins deux classes d'équivalence. Soit alors $\mathcal{B}'$ une autre base. On a : $P_{\mathcal{B},\mathcal{B}'} = P_{\mathcal{B}^{\mathrm{op}},\mathcal{B}'} P_{\mathcal{B},\mathcal{B}^{\mathrm{op}}}$, et $\det P_{\mathcal{B},\mathcal{B}^{\mathrm{op}}} < 0$. Par suite, on peut affirmer que $\det P_{\mathcal{B},\mathcal{B}'}$ ou $\det P_{\mathcal{B},\mathcal{B}^{\mathrm{op}}}$ est positif, donc que $\mathcal{B}'$ définit la même orientation que $\mathcal{B}$ ou que $\mathcal{B}^{\mathrm{op}}$.□

Remarque. **Orientation induite dans un hyperplan.** Soit E un espace de dimension finie et H un hyperplan. La donnée d'un vecteur e hors de H induit une orientation de H : une base $(e_1, \ldots, e_{n-1})$ de H est directe si et seulement si la base $(e, e_1, \ldots, e_{n-1})$ de E est directe. Inversement, la donnée d'une orientation sur H ne permet pas de définir une orientation sur E.

Exemple : Soit un plan orienté. De façon suggestive, on représente une orientation par un "arc de cercle orienté". Fixons une droite H. Sans plus de données, rien à faire pour orienter H. Mais si on se donne un vecteur e hors de cette droite, il n'y a plus qu'une seule orientation de H compatible avec celle du plan au sens précédent : sur la figure, c'est l'orientation donnée par e_1, alors que celle donnée par e_1' ne convient pas. Evident, non ?



Idem avec un plan dans l'espace (penser au célèbre "bonhomme d'Ampère").

Chapitre 4
Réduction des endomorphismes

Idée : Etant donné un endomorphime d'un espace vectoriel, trouver des bases dans lesquelles la matrice soit aussi simple que possible.

Notations : On fixe donc un espace vectoriel E de dimension n sur un corps $\mathbb{K}$, un endomorphisme φ de E et une matrice $n \times n$ notée A. On note alors φ_A l'endomorphisme de $\mathbb{K}^n$ défini par : $\varphi_A(X) = AX$ pour $X \in \mathbb{K}^n$.

Vocabulaire : voir paragraphe II.

I Préliminaires

1° Polynômes d'endomorphismes

a) On définit φ^k (resp. A^k) par récurrence en posant : $\varphi^0 = \mathrm{Id}$, $\varphi^{k+1} = \varphi^k \circ \varphi = \varphi \circ \varphi^k$ (resp. $A^0 = I_n$, $A^{k+1} = A^k A = AA^k$). Pour $P \in \mathbb{K}[X]$, $P = \sum_{k=0}^{d} a_k X^k$, on pose alors

$$P(\varphi) = \sum_{k=0}^{d} a_k \varphi^k \quad (\text{resp. } P(A) = \sum_{k=0}^{d} a_k A^k).$$

Lemme *Les applications* $\mathrm{eval}_\varphi : \mathbb{K}[X] \to \mathcal{L}(E)$, $P \mapsto P(\varphi)$ *et* $\mathrm{eval}_A : \mathbb{K}[X] \to \mathcal{M}_n(\mathbb{K})$, $P \mapsto P(A)$ *sont des morphismes d'algèbre.*

Attention ! *La loi sur* $\mathbb{K}[X]$ *que l'on considère est le produit des polynômes, pas la composition (ou substitution) qui d'ailleurs ne fait pas de* $\mathbb{K}[X]$ *une algèbre (cette loi n'est pas distributive sur* $+$*).*

Remarque. On notera $PQ(\varphi)$ au lieu de $(PQ)(\varphi)$. La seule chose à démontrer pour le lemme est que $PQ(\varphi) = P(\varphi) \circ Q(\varphi)$. Notons que le lemme entraîne que $P(\varphi) \circ Q(\varphi) = PQ(\varphi) = QP(\varphi) = Q(\varphi) \circ P(\varphi)$ (idem avec A).

b) Polynôme annulateur. On dit que $P \in \mathbb{K}[X]$ est un polynôme annulateur de φ (resp. de A) si $P(\varphi) = 0 \in \mathcal{L}(E)$ (resp. si $P(A) = 0 \in \mathcal{M}_n(\mathbb{K})$).

Remarque. P annule φ si et seulement si P annule la matrice de φ dans une base donnée si et seulement si P annule la matrice de φ dans n'importe quelle base.

c) Complément hors programme : polynôme minimal. Considérons le noyau du morphisme d'évaluation eval_φ (ou eval_A) : c'est un idéal de $\mathbb{K}[X]$, donc il est engendré par un polynôme, qu'on rend unique en le choisissant unitaire. On le note μ_φ (ou μ_A) et on l'appelle le polynôme minimal de φ (ou A). Il est caractérisé par la propriété de diviser tous les polynômes qui annulent φ (ou A). (Détails sur les feuilles d'exercices.)

2° Lemme des noyaux

Proposition *Soit* $P, Q \in \mathbb{K}[X]$ *deux polynômes premiers entre eux. Alors on a :*

$$\mathrm{Ker}\, PQ(\varphi) = \mathrm{Ker}\, P(\varphi) \oplus \mathrm{Ker}\, Q(\varphi) \quad (\textit{resp.}\, \mathrm{Ker}\, PQ(A) = \mathrm{Ker}\, P(A) \oplus \mathrm{Ker}\, Q(A)).$$

Démonstration. Le théorème de Bezout assure l'existence de $U, V \in \mathbb{K}[X]$ tels que $UP + VQ = 1 \in \mathbb{K}[X]$. On a donc : $VQ(\varphi) + UP(\varphi) = \mathrm{Id}$ ou, pour tout $v \in E$:

$$(\S) \qquad v = VQ(\varphi)(v) + UP(\varphi)(v).$$

Intérêt : si $v \in \operatorname{Ker} PQ(\varphi)$, alors : $VQ(\varphi)(v) \in \operatorname{Ker} P(\varphi)$ et $UP(\varphi)(v) \in \operatorname{Ker} Q(\varphi)$. (En effet, on a par exemple : $P(\varphi)[VQ(\varphi)(v)] = PVQ(\varphi)(v) = V(\varphi)[PQ(\varphi)(v)] = 0$.) Ainsi, (§) exprime v comme somme d'un vecteur de $\operatorname{Ker} P(\varphi)$ et d'un vecteur de $\operatorname{Ker} Q(\varphi)$.
Soit $v \in \operatorname{Ker} P(\varphi) \cap \operatorname{Ker} Q(\varphi)$. On a : $P(\varphi)(v) = 0$ donc $UP(\varphi)(v) = U(\varphi)[P(\varphi)(v)] = 0$ et de même, $VQ(\varphi)(v) = 0$. Par (§) il vient : $v = 0$, ce qui montre que la somme est directe.□

Application : Le théorème des accroissements finis permet de montrer que les seules fonctions dérivables dont la dérivée est nulle sont les constantes. On en déduit que les vecteurs propres de l'opérateur $D : y \mapsto y'$ sur l'espace des fonctions C^∞ sont les multiples des fonctions exponentielles $x \mapsto \exp(\lambda\, x)$.
Notons que la démonstration du lemme s'applique même si E est de dimension infinie. Le lemme des noyaux permet alors de résoudre les équations différentielles linéaires du genre : $P(D)(y) = 0$, lorsque P est un polynôme sans racine multiple. Ex : $y'' - 5y' + 6y = 0$.

3° Un résultat d'indépendance linéaire

Proposition *Des vecteurs propres de φ associées à des valeurs propres distinctes sont linéairement indépendants.*

Démonstration. Soit $v_1, \ldots, v_r$ des vecteurs propres et $\lambda_1, \ldots, \lambda_r$ les valeurs propres associées, que l'on suppose distinctes. Supposons que ces vecteurs soient liés. Parmi l'ensemble des combinaisons linéaires nulles à coefficients non tous nuls[1] des v_i, choisissons-en une qui fasse intervenir un nombre minimal s de vecteurs. En d'autres termes, on choisit une sous-famille liée dont toute sous-famille stricte est libre. Quitte à renuméroter les vecteurs et à en supprimer si nécessaire, on peut supposer que $s = r$.
Ainsi, on a r scalaires $\alpha_1, \ldots, \alpha_r$ tous non nuls[2] tels que $\sum_{i=1}^r \alpha_i\, v_i = 0$, et toute sous-famille stricte de $(v_1, \ldots, v_r)$ est libre. Appliquons φ d'une part, multiplions par λ_r d'autre part, et soustrayons membre à membre : il vient $\sum_{i=1}^{r-1} \alpha_i(\lambda_i - \lambda_r)\, v_i = 0$. Par indépendance linéaire de la famille $(v_1, \ldots, v_{r-1})$, on a : $\alpha_i(\lambda_i - \lambda_r) = 0$, ce qui donne, puisque $\lambda_i \neq \lambda_r$: $\alpha_i = 0$ (pour $i \leq r-1$). On reporte dans la combinaison linéaire initiale pour trouver : $\alpha_r\, v_r = 0$, ce qui est contradictoire.□

4° Sous-espaces stables

a) On dit que $F \subset E$ est stable par φ si $\varphi(F) \subset F$, i.e. si pour tout $v \in F$, on a $\varphi(v) \in F$.

Attention ! *Un sous-espace stable n'a pas toujours de supplémentaire stable. (Ceci entraînerait que toute matrice complexe est diagonalisable –pourquoi ?) Par exemple, prendre en dimension 2, $\varphi(e_1) = e_1$ et $\varphi(e_2) = e_1 + e_2$, et $F = \operatorname{Vect}(e_1)$.*

Un intérêt des sous-espaces stables. Si F est stable par φ, choisissons une base de F et complétons-la en une base de E. Alors la matrice de φ dans cette base est triangulaire par blocs. De plus, le bloc en haut à gauche est la matrice de la restriction de φ à F.
Si F et G sont deux supplémentaires stables, on forme une base de E en juxtaposant une base de F et une base de G. La matrice de φ dans cette base est alors diagonale par blocs.

Lemme *Si $\varphi \circ \psi = \psi \circ \varphi$ alors $\operatorname{Im} \psi$ et $\operatorname{Ker} \psi$ sont des sous-espaces stables par φ.*

Exemples : $\psi = \varphi$, $\psi = (\varphi - \lambda \mathrm{Id})^k$,...

[1] Distinguer entre "non tous nuls" et "tous non nuls."
[2] Idem.

II Vocabulaire

1° Eléments propres

a) On appelle *valeur propre* de l'endomorphisme φ tout scalaire $\lambda \in \mathbb{K}$ satisfaisant les conditions équivalentes suivantes :

- $\mathrm{Ker}(\varphi - \lambda \mathrm{Id}) \neq \{0\}$;
- il existe $v \in E$, $v \neq 0$, tel que $\varphi(v) = \lambda\, v$;
- $\det(\varphi - \lambda \mathrm{Id}) \neq 0$.

On appelle *valeur propre* de la matrice A une valeur propre de l'endomorphisme φ_A. L'ensemble des valeurs propres d'un endomorphisme ou d'une matrice est son *spectre*.

b) On appelle *vecteur propre* de φ associé à la valeur propre λ tout vecteur non nul de $\mathrm{Ker}(\varphi - \lambda \mathrm{Id})$, c'est-à-dire un vecteur non nul v tel que $\varphi(v) = \lambda\, v$.

On appelle *vecteur propre* de A tout vecteur propre de φ_A.

c) On appelle *espace propre* de φ associé à la valeur propre λ le sous-espace $\mathrm{Ker}(\varphi - \lambda \mathrm{Id})$.

On appelle *espace propre* de A associé à la valeur propre λ le sous-espace $\mathrm{Ker}(\varphi_A - \lambda \mathrm{Id})$.

2° Polynôme caractéristique

a) On appelle *polynôme caractéristique* de φ (resp. de A) le polynôme $\chi_\varphi \in \mathbb{K}[X]$ (resp. $\chi_A \in \mathbb{K}[X]$) défini par :

$$\chi_\varphi(X) = \det(\varphi - X\,\mathrm{Id}) \qquad (\text{resp. } \chi_A(X) = \det(A - X\,\mathrm{Id})).$$

Sens : En principe, on ne sait définir que les déterminants de matrices sur un corps, alors que là, les coefficients sont dans l'algèbre $\mathbb{K}[X]$, qui n'est pas un corps. On s'en tire en passant dans le corps des fractions rationnelles $\mathbb{K}(X)$ de la façon suivante :

Soit $\mathcal{B}$ une base de E, et A la matrice de φ dans $\mathcal{B}$. Alors on pose : $\chi_\varphi(X) = \det(A - X\mathrm{Id}) \in \mathbb{K}(X)$. Vérifions que ceci a un sens.

D'une part, soit $\mathcal{B}'$ est une autre base de E, P la matrice de passage, A' la matrice de φ dans $\mathcal{B}'$. On a $A' = PAP^{-1}$ donc $A' - X\mathrm{Id} = P(A - X\mathrm{Id})P^{-1}$ (dans les matrices à coefficients dans le corps $\mathbb{K}(X)$), d'où : $\det(A' - X\mathrm{Id}) = \det(P)\det(A - X\mathrm{Id})\det(P)^{-1}$ dans $\mathbb{K}(X)$. Ceci montre que $\chi_\varphi(X)$ ne dépend que de φ, et pas du choix de $\mathcal{B}$.

D'autre part, on constate grâce à la formule qui le définit que $\det(A - X\mathrm{Id})$ est bien un polynôme en X (cela aurait pu être une fraction rationnelle non polynômiale).

Remarque : Le polynôme caractéristique est de la forme :

$$\chi_\varphi(X) = (-1)^n \left[X^n - \mathrm{tr}(\varphi)\, X^{n-1} + \cdots + (-1)^n \det(\varphi)\right]$$

$$(\text{resp.} \chi_A(X) = (-1)^n \left[X^n - \mathrm{tr}(A)\, X^{n-1} + \cdots + (-1)^n \det(A)\right]).$$

b) Multiplicité des valeurs propres.

Proposition *Notons n_λ la multiplicité de λ comme racine de χ_φ. Alors on a :*

$$\dim \mathrm{Ker}(\varphi - \lambda\,\mathrm{Id}) \leq n_\lambda.$$

Démonstration. Le sous-espace $E_\lambda = \mathrm{Ker}(\varphi - \lambda\,\mathrm{Id})$ est stable par φ. La restriction $\varphi|_{E_\lambda}$ de φ à E_λ est une homothétie donc son polynôme caractéristique est $(X - \lambda)^{\dim E_\lambda}$. Fixons une base de E_λ et complétons-la en une base de E. Si on calcule le polynôme caractéristique dans cette base, on voit que le polynôme caractéristique de $\varphi|_{E_\lambda}$ divise celui de φ, ce qui entraîne l'assertion.$\square$

c) Théorème de Cayley–Hamilton : $\chi_\varphi(\varphi) = 0$ (resp. $\chi_A(A) = 0$).
Démonstration. Voir plus loin.
Exemple 1 : Pour $A \in \mathcal{M}_2(\mathbb{K})$, on vérifie que $\chi_A(X) = X^2 - \mathrm{tr}(A)\, X + \det(A)$ et le théorème de Cayley-Hamilton est facile.
Exemple 2 : Fixons $P = X^n + \sum_{k=1}^{n-1} a_k X^k \in \mathbb{K}[X]$ et posons :

$$A = A_P = \begin{pmatrix} 0 & \dots & \dots & 0 & -a_0 \\ 1 & \ddots & & \vdots & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & \dots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

C'est la matrice-compagnon de P. On peut calculer à la main : $\chi_A(X) = (-1)^n P(X)$. Dans ce cas, on vérifie le théorème de Cayley-Hamilton en montrant que $A^k e_1 = e_{k+1}$ (pour $1 \leq k \leq n-1$), puis que $P(A)(e_1) = 0$ et enfin que $P(A)(e_k) = 0$.

Remarque. Ce raisonnement permet de montrer facilement que $\mu_A = P$.
d) Intermède technique.

Lemme *Si P annule φ (resp. A), alors toute racine de χ_φ est une racine de P.*

Démonstration. Soit λ une racine de χ_φ et soit v un vecteur propre. Alors, on a dans E : $0 = P(\varphi)(v) = P(\lambda)\, v$ donc $P(\lambda) = 0$, puisque $v \neq 0$.□

Remarque. Inversement, par définition, μ_φ divise χ_φ donc toute racine de μ_φ est une racine de χ_φ. Ainsi, le polynôme minimal et le polynôme caractéristique ont exactement les mêmes racines (mais, en général, avec des multiplicités différentes).

III Diagonalisation

1° Définition

On dit que φ est *diagonalisable* si les conditions équivalentes suivantes sont vérifiées :
- E est la somme directe des espaces propres de φ ;
- E possède une base formée de vecteurs propres de φ ;
- E possède une base dans laquelle la matrice de φ est diagonale.

Exemple 1 : toute matrice diagonale est diagonalisable. (Pourquoi ?)
Exemple 2 : la matrice $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ est diagonalisable si et seulement si $a = 0$. (Pourquoi ?)
Exemple 3 : L'exemple 2 se généralise ainsi : si une matrice est triangulaire et si tous les coefficients diagonaux sont égaux, alors elle est diagonalisable si et seulement si c'est une homothétie. Attention, ceci n'est valable que si les coefficients diagonaux sont tous égaux !

2° Critères de diagonalisabilité.

Proposition *Les conditions suivantes sont équivalentes :*
(i) *φ est diagonalisable ;*
(ii) *φ est annulé par un polynôme scindé à racines simples ;*
(iii) *χ_φ est scindé et pour toute racine λ de χ_φ de multiplicité n_λ, on a :* $\dim \mathrm{Ker}(\varphi - \lambda \mathrm{Id}) = n_\lambda$;
(iv) *le polynôme minimal de φ est scindé à racines simples (chut, ne l'ébruitez pas !).*

Corollaire *Si χ_φ (resp. χ_A) est scindé à racines simples, alors φ (resp. A) est diagonalisable.*

Exemples classiques : $\varphi^2 = \varphi$ (projecteurs), $\varphi^2 = \mathrm{Id}$ (symétries), etc.
Démonstration. (i)⇔(ii) : si φ est diagonalisable, alors φ annule le polynôme $P(X) = \prod_{\lambda \in (\varphi)} (X - \lambda)$, où (φ) est le spectre de φ. La réciproque résulte du lemme des noyaux.

(i)$\Leftrightarrow$(iii) : Sens direct : calculer le polynôme caractéristique dans une base qui diagonalise. Réciproquement, en juxtaposant des bases des divers $\mathrm{Ker}(\varphi - \lambda \mathrm{Id})$, on obtient une famille libre de cardinal $\sum n_\lambda = \deg \chi_\varphi = \dim E$.$\square$

IV Trigonalisation

1° Cas des matrices nilpotentes

a) On dit que φ (resp. A) est nilpotent si $\varphi^k = 0$ (resp. $A^k = 0$) pour un certain $k \in \mathbb{N}^*$.

Proposition *Si φ est nilpotent, il existe une base de E dans laquelle la matrice de φ est triangulaire stricte. Si A est nilpotente, elle est semblable à une matrice triangulaire stricte.*

Démonstration. On note que pour tout i, on a :

$$\mathrm{Ker}\,\varphi \subset \mathrm{Ker}\,\varphi^2 \subset \cdots \subset \mathrm{Ker}\,\varphi^i \subset \mathrm{Ker}\,\varphi^{i+1} \subset \cdots \subset \mathrm{Ker}\,\varphi^k = E.$$

On fixe une base $(e_1, \ldots, e_{d_1})$ de $\mathrm{Ker}\,\varphi$, que l'on complète en une base $(e_1, \ldots, e_{d_1+d_2})$ de $\mathrm{Ker}\,\varphi^2$, [...] que l'on complète en une base $(e_1, \ldots, e_n)$ de $\mathrm{Ker}\,\varphi^k = E$.
En remarquant que pour $v \in \mathrm{Ker}\,\varphi^i$, on a : $\varphi(v) \in \mathrm{Ker}\,\varphi^{i-1}$, on obtient que la matrice de φ dans cette base est triangulaire par blocs (les blocs diagonaux ayant pour taille $d_i \times d_i$).$\square$

Conséquence : le polynôme caractéristique de φ (ou A) est $(-1)^n X^n$ et $\varphi^n = 0$ (ou $A^n = 0$).

Remarque. On peut mieux faire avec plus de travail (voir la forme de Jordan ci-dessous.)

2° Théorème de trigonalisation

Théorème *Si χ_φ est scindé, alors il existe une base dans laquelle la matrice de φ est triangulaire supérieure. Si χ_A est scindé, alors A est semblable à une matrice triangulaire.*

Cas particulier important : lorsque $\mathbb{K}$ est algébriquement clos, tous les polynômes sont scindés. Par exemple, toute matrice carrée complexe est trigonalisable.
Démonstration. Par récurrence sur $n = \dim E$. Pour $n = 1$, la propriété est triviale. Supposons-la démontrée jusqu'au rang $n-1$. Soit donc A une matrice $n \times n$.
Par hypothèse, χ_A possède une racine λ. Soit $v \in \mathbb{K}^n$ un vecteur propre associé, qu'on complète en une base $\mathcal{B}'$ de $\mathbb{K}^n$. La matrice de passage P de la base canonique à la base $\mathcal{B}'$ montre que A est semblable à une matrice triangulaire par blocs :

$$P^{-1}AP = \left(\begin{array}{c|c} \lambda & C \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right).$$

On a alors : $\chi_A(X) = \chi_B(X)\,(X - \lambda)$, donc χ_B est scindé, donc on peut trouver une matrice inversible Q de taille $(n-1) \times (n-1)$ telle que $Q^{-1}BQ = T'$ soit triangulaire supérieure. On a alors :

$$\left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & Q \\ 0 & \end{array}\right)^{-1} \left(\begin{array}{c|c} \lambda & C \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right) \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & Q \\ 0 & \end{array}\right) = \left(\begin{array}{c|c} \lambda & C' \\ \hline 0 & \\ \vdots & T' \\ 0 & \end{array}\right).$$

Cette dernière matrice est semblable à A et triangulaire, donc on a gagné.$\square$

Corollaire (Cayley-Hamilton) $\chi_\varphi(\varphi) = 0 \in \mathcal{L}(E)$, $\chi_A(A) = 0 \in \mathcal{M}_n(\mathbb{K})$.

Démonstration. On choisit une base $(e_1, \ldots, e_n)$ de E dans laquelle la matrice de φ est triangulaire. Appelons $\lambda_1, \ldots, \lambda_n$ les coefficients diagonaux. On note alors : $F_i = \mathrm{Vect}(e_1, \ldots, e_i)$. Il est clair que F_i est stable par φ, que la restriction φ_i de φ à F_i a pour polynôme caractéristique $\chi_i(X) = \prod_{k=1}^{i}(X-\lambda_k)$. On montre alors par une récurrence pas si compliquée que : $\chi_i(\varphi_i) = 0$ (appliquer à $e_1, \ldots, e_{i-1}$ puis à e_i). Lorsque $i = n$, on a gagné.
Euh, il y a une escroquerie, ici : laquelle ? Comment peut-on y remédier ?□

3° Sous-espaces caractéristiques

Factorisons χ_φ sur $\mathbb{K}$. En notant $\lambda_1, \ldots, \lambda_r$ les racines dans $\mathbb{K}$ et $n_1, \ldots, n_r$ leurs multiplicités respectives, on fait apparaître un polynôme $Q \in \mathbb{K}[X]$ sans racine dans $\mathbb{K}$ tel que :

$$\chi_\varphi(X) = \prod_{i=1}^{r}(X-\lambda_i)^{n_i}\, Q(X).$$

D'après le théorème de Cayley-Hamilton et le lemme des noyaux, on a donc :

$$E = \bigoplus_{i=1}^{r} \mathrm{Ker}\,((\varphi - \lambda_i \mathrm{Id})^{n_i}) \oplus \mathrm{Ker}\, Q(\varphi).$$

Les sous-espaces $\mathrm{Ker}\,((\varphi - \lambda_i \mathrm{Id})^{n_i})$ sont appelés sous-espaces caractéristiques de φ.

Proposition $\dim \mathrm{Ker}\,((\varphi - \lambda_i \mathrm{Id})^{n_i}) = n_i$.

Démonstration. Notons $E^{(\lambda_i)} = \mathrm{Ker}\,((\varphi - \lambda_i \mathrm{Id})^{n_i})$ et $d_i = \dim E^{(\lambda_i)}$. Par définition, la restriction de $\varphi - \lambda_i \mathrm{Id}$ à $E^{(\lambda_i)}$ est nilpotente, donc son polynôme caractéristique est $(-1)^{d_i} X^{d_i}$, donc celui de la restriction de φ à $E^{(\lambda_i)}$ est $(X-\lambda_i)^{d_i}$.
En utilisant des bases des divers $E^{(\lambda_i)}$ et de $\mathrm{Ker}\, Q(\varphi)$ pour calculer χ_φ, on se convainc que

$$(\spadesuit) \qquad \prod_{i=1}^{r}(X-\lambda_i)^{n_i}\, Q(X) = \chi_\varphi(X) = \prod_{i=1}^{r}(-1)^{d_i}(X-\lambda_i)^{d_i}.\tilde{Q}(X),$$

où $\tilde{Q}$ est le polynôme caractéristique de la restriction de φ à $\mathrm{Ker}\, Q(\varphi)$. On est sûr que $\tilde{Q}(\lambda_i) \neq 0$, sans quoi λ_i est une valeur propre de la restriction de φ à $\mathrm{Ker}\, Q(\varphi)$, et le vecteur propre correspondant appartiendrait à $E^{(\lambda_i)} \cap \mathrm{Ker}\, Q(\varphi)$. En identifiant les facteurs premiers de l'égalité $(\spadesuit)$, on obtient : $n_i = d_i$.□
Pour la suite, il est commode de montrer le

Lemme *Reprenons les notations précédentes et supposons que χ_φ soit scindé, i.e. que $Q = 1$. Alors, pour tout i, le projecteur π_i d'image $\mathrm{Ker}(\varphi - \lambda_i \mathrm{Id})^{n_i}$ et de noyau $\bigoplus_{j \neq i} \mathrm{Ker}(\varphi - \lambda_j \mathrm{Id})^{n_j}$ est un polynôme en φ. De plus, on a : $\sum_{i=1}^{r} \pi_i = \mathrm{Id}$.*

Démonstration. Notons $P_i(X) = (X-\lambda_i)^{n_i}$ et $Q_i = \prod_{j \neq i}(X-\lambda_j)^{n_j}$. Comme P_i et Q_i sont premiers entre eux, il existe $U_i, V_i \in \mathbb{K}[X]$ tels que $V_i Q_i + U_i P_i = 1$. Il est facile de vérifier que $\pi_i = V_i Q_i(\varphi)$ et que $\sum_{i=1}^{r} \pi_i = \mathrm{Id}$.□

4° Décomposition de Jordan

Théorème *Soit E de dimension finie et $\varphi \in \mathcal{L}(E)$. Supposons que le polynôme caractéristique de φ soit scindé (c'est par exemple le cas si $\mathbb{K}$ est algébriquement clos).*
Alors il existe un unique couple $(\delta, \eta) \in \mathcal{L}(E) \times \mathcal{L}(E)$ tel que :
(i) $\varphi = \delta + \eta$;
(ii) *δ est diagonalisable et η nilpotente ;*
(iii) $\varphi \circ \delta = \delta \circ \varphi$ *et* $\varphi \circ \eta = \eta \circ \varphi$.
De plus, δ et η sont alors des polynômes en φ.

Démonstration. 1^e étape : Existence si φ a une unique valeur propre. On pose alors $\delta = \lambda\mathrm{Id}$ et $\eta = \varphi - \lambda\mathrm{Id}$. Le théorème de Cayley-Hamilton montre que δ et η conviennent.
L'idée consiste alors à se ramener à ce cas particulier.
2^e étape : Existence dans le cas général. On reprend les notations du lemme 3°. On pose :

$$\delta = \sum_{i=1}^{r} \lambda_i \, \pi_i \text{ et } \eta = \sum_{i=1}^{r} (\varphi - \lambda_i) \, \pi_i.$$

Montrons que δ et η conviennent.
(i) $\delta + \eta = \sum_{i=1}^{r} \varphi \, \pi_i = \varphi \sum_{i=1}^{r} \pi_i = \varphi$ car $\sum_{i=1}^{r} \pi_i = \mathrm{Id}$.
(ii) Les espaces caractéristiques $\mathrm{Ker}(\varphi - \lambda_i \mathrm{Id})^{n_i}$ sont supplémentaires et stables par δ et η. La restriction de δ à chacun est $\lambda_i \mathrm{Id}$, i.e. une homothétie, donc δ est diagonalisable. La restriction de η à chacun est $(\varphi - \lambda_i \mathrm{Id})$, donc nilpotente, donc η est nilpotente.
(iii) δ et η sont des polynômes en φ donc commutent à φ.

3^e étape : Unicité. Notons (δ_0, η_0) le couple obtenu dans la 2^e étape et soit (δ, η) un autre couple vérifiant (i), (ii) et (iii). On a alors par (iii), vu que δ_0 et η_0 sont des polynômes en φ : $\delta \, \delta_0 = \delta_0 \, \delta$ et $\eta \, \eta_0 = \eta_0 \, \eta$. Par suite, $\delta - \delta_0$ est diagonalisable (voir exercices) et $\eta_0 - \eta$ est nilpotente (par la formule du binôme). Or, on a : $\varphi = \delta + \eta = \delta_0 + \eta_0$ donc $\delta - \delta_0 = \eta_0 - \eta$. Et on sait que la seule application diagonalisable et nilpotente est l'application nulle. D'où $\delta = \delta_0$ et $\eta = \eta_0$.□

5° Décomposition de Jordan (une autre, hors programme)

Théorème *Soit E de dimension finie et $\varphi \in \mathcal{L}(E)$. Supposons que le polynôme caractéristique de φ soit scindé (c'est par exemple le cas si $\mathbb{K}$ est algébriquement clos). Alors dans une base convenable, la matrice de φ est diagonale par blocs, chaque bloc étant une matrice $r \times r$ de la forme :*

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}.$$

De plus, le nombre de blocs de taille donnée ayant une valeur propre donnée sont déterminés de façon unique par φ.

Remarque : Si $r = 1$, le bloc correspondant est juste (λ). Si une matrice est diagonalisable, tous les blocs sont de ce type.

Exercice : Calculer le polynôme caractéristique d'une matrice de la forme précédente. En déduire la forme de Jordan d'une matrice compagnon.

Exercice : Déduire du théorème précédent que toute matrice est semblable à sa transposée. (Commencer par le cas où il n'y a qu'un seul bloc.)

Chapitre 5
Espaces euclidiens et hermitiens

I Généralités

1° Espaces euclidiens et hermitiens

a) Définitions

Un espace *euclidien* est un espace vectoriel E de dimension finie sur $\mathbb{R}$, muni d'un produit scalaire, i.e. d'une forme $\langle .,. \rangle : E \times E \to \mathbb{R}$ bilinéaire-symétrique-définie-positive. Cela signifie que pour $v, v', w \in E$ et $\lambda, \lambda' \in \mathbb{R}$:

- $\langle \lambda v + \lambda' v', w \rangle = \lambda \langle v, w \rangle + \lambda' \langle v', w \rangle$ et $\langle w, \lambda v + \lambda' v' \rangle = \lambda \langle w, v \rangle + \lambda' \langle w, v' \rangle$;
- $\langle v, w \rangle = \langle w, v \rangle$;
- $\langle v, v \rangle \geq 0$ et si $\langle v, v \rangle = 0$, alors $v = 0$.

Un espace *hermitien* est un espace vectoriel E de dimension finie sur $\mathbb{C}$, muni d'un produit scalaire, i.e. d'une forme $\langle .,. \rangle : E \times E \to \mathbb{R}$ sesquilinéaire-à symétrie hermitienne-définie-positive. Cela signifie que pour $v, v', w \in E$ et $\lambda, \lambda' \in \mathbb{R}$:

- $\langle \lambda v + \lambda' v', w \rangle = \overline{\lambda} \langle v, w \rangle + \overline{\lambda'} \langle v', w \rangle$ et $\langle w, \lambda v + \lambda' v' \rangle = \lambda \langle w, v \rangle + \lambda' \langle w, v' \rangle$;
- $\langle v, w \rangle = \overline{\langle w, v \rangle}$;
- $\langle v, v \rangle \geq 0$ et si $\langle v, v \rangle = 0$, alors $v = 0$.

Désormais, E est un espace euclidien ou hermitien.

Notation traditionnelle : Pour $v \in E$, on note : $||v|| = \sqrt{\langle v, v \rangle}$.
Vocabulaire : Si on abandonne l'hypothèse de la dimension finie, on parle d'espace pré-hilbertien (réel ou complexe). Si on ajoute alors l'hypothèse que l'espace est complet, on parle d'espace hilbertien.

Exemple : Etant donné une base $(e_1, \ldots, e_n)$ d'un espace vectoriel E de dimension finie sur $\mathbb{R}$ ou $\mathbb{C}$, il existe un unique produit scalaire tel que $\langle e_i, e_j \rangle = \delta_{ij}$ ($i, j = 1, \ldots, n$, δ est le δ de Kronecker). Si $E = \mathbb{R}^n$ ou $\mathbb{C}^n$, le produit scalaire canonique/standard est le produit scalaire construit ainsi à partir de la base canonique/standard.

b) Produit scalaire et dualité

Ici, pour simplifier, on suppose que E est euclidien. Comme toute forme bilinéaire sur E, un produit scalaire permet de définir une application linéaire I de E vers son dual E^* par : $I : E \to E^*$, $v \mapsto I_v$, où $I_v : E \to \mathbb{R}$, $w \mapsto I_v(w) = \langle v, w \rangle$. Mais ici, on peut affirmer que I est un isomorphisme, car le noyau de I est trivial (si $I_v = 0$, alors $I_v(v) = 0$ donc $v = 0$), et on a supposé E de dimension finie.
Ainsi, le choix d'un produit scalaire permet d'identifier un espace euclidien et son dual.

c) Orthogonal d'une partie

On définit naturellement l'orthogonal d'une partie $P \subset E$ (au sens du produit scalaire) comme l'ensemble des vecteurs $v \in E$ tels que $\langle v, w \rangle = 0$ pour tout $w \in P$. La relation suivante est immédiate : $P^\perp = \mathrm{Vect}(P)^\perp$.
Soit F un sous-espace de E. Soit $P = (e_1, \ldots, e_r)$ une base quelconque de F. La remarque précédente montre que $F^\perp$ est l'intersection des noyaux des r formes linéaires $v \mapsto \langle e_i, v \rangle$ ($i = 1, \ldots, r$), lesquelles sont linéairement indépendantes (vérifier). Par suite, on a : $\dim F^\perp = \dim E - \dim F$. On note que $F \cap F^\perp = \{0\}$ car $\langle .,. \rangle$ est définie positive, ce qui entraîne :

Lemme *L'orthogonal d'un sous-espace est un supplémentaire de ce sous-espace.*

Ainsi, en présence d'un produit scalaire, un sous-espace possède un supplémentaire naturel (défini sans choix supplémentaire), qui est son orthogonal. Notons que l'orthogonal dépend quand même du choix du produit scalaire...

2° Bases orthonormées

a) Définition et formules

Une base $(e_1, \ldots, e_n)$ de E est dite orthonormée si pour tout $i, j = 1, \ldots, n$, on a : $\langle e_i, e_j \rangle = \delta_{ij}$ (le δ de Kronecker). Les bases orthonormées permettent des calculs commodes :

Lemme *Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée.*
(i) *Pour $v \in E$, les coordonnées de v dans $\mathcal{B}$ sont données par :*

$$v = \sum_{i=1}^{n} \langle e_i, v \rangle \, e_i.$$

(ii) *Pour $\varphi \in \mathcal{L}(E)$, la matrice $A = (a_{ij})_{i,j=1,\ldots,n}$ de φ dans $\mathcal{B}$ est donnée par :*

$$a_{ij} = \langle e_i, \varphi(e_j) \rangle.$$

On peut montrer l'existence de bases orthonormées par une récurrence facile sur la dimension. Mais on peut faire mieux :

b) Existence de bases orthonormées

Proposition (Gram-Schmidt) *Si $(v_1, \ldots, v_n)$ est une base de E, il existe une base orthonormée $(e_1, \ldots, e_n)$ de E (unique à changement de signe des vecteurs près) telle que pour $i = 1, \ldots, n$, on ait :* $\mathrm{Vect}(e_1, \ldots, e_i) = \mathrm{Vect}(v_1, \ldots, v_i)$.

Démonstration. On prend $e_1 = v_1/||v_1||$. Supposons avoir construit $e_1, \ldots, e_i$. On cherche e'_{i+1} sous la forme $v_{i+1} + \sum_{k=1}^{i} \lambda_k e_k$ de sorte que e_{i+1} soit orthogonal à $e_1, \ldots, e_i$. Ceci force à prendre $\lambda_k = -\langle v_{i+1}, e_k \rangle$. Enfin, on pose $e_{i+1} = e'_{i+1}/||e'_{i+1}||$. On vérifie facilement que cela convient et qu'il n'y a pas d'autre choix (aux signes près).$\square$

Exercice. En déduire que toute matrice s'écrit de façon essentiellement unique comme produit d'une matrice orthogonale (voir ci-dessous) et d'une matrice triangulaire supérieure. Cela s'appelle la décomposition d'Iwasawa.

c) Projecteurs orthogonaux

Un projecteur est dit orthogonal si son image et son noyau sont orthogonaux. Si $\pi : E \to E$ en est un, si $(e_1, \ldots, e_r)$ est une base orthonormée de $\mathrm{Im}\,\pi$, on a pour $v \in E$ (vérifier !) :

$$\pi(v) = \sum_{i=1}^{r} \langle e_i, v \rangle \, e_i.$$

D'un autre point de vue, si F est un sous-espace de E, si $(f_1, \ldots, f_d)$ est une base orthonormée de F, on constate que pour tout $v \in E$, le vecteur $\pi(v) = v - \sum_{i=1}^{d} \langle d_i, v \rangle \, d_i$ est orthogonal à F. Ceci prouve à nouveau que F et $F^\perp$ engendrent E, et, en notant que $F \cap F^\perp = \{0\}$, qu'ils sont supplémentaires. De plus, π est la projection sur $F^\perp$ parallèlement à F.

II Endomorphismes auto-adjoints et unitaires

1° Adjoint

Proposition (i) *Pour $\varphi \in \mathcal{L}(E)$, il existe un unique $\varphi^* \in \mathcal{L}(E)$ tel que*

$$\forall v, w \in E, \quad \langle v, \varphi(w) \rangle = \langle \varphi^*(v), w \rangle.$$

(ii) *La matrice de φ^* dans une base orthonormée est la conjuguée (coefficient par coefficient) de la transposée de la matrice de φ.*

Définitions : Avec les notations ci-dessus, on dit que φ^* est l'adjoint de φ. Pour une matrice $A \in \mathcal{M}_n(\mathbb{C})$, la conjuguée de sa transposée, $A^* = {}^t\overline{A}$, est l'adjointe de A.

Démonstration. Fixons une base orthonormée $\mathcal{B} = (e_1, \ldots, e_n)$. Si φ^* existe, on doit avoir : $\langle \varphi^*(e_i), e_j \rangle = \langle e_i, \varphi(e_j) \rangle$, donc la matrice de φ^* dans $\mathcal{B}$ est nécessairement la conjuguée de la transposée de celle de φ. Inversement, l'endomorphisme qui a pour matrice la conjuguée de la transposée de la matrice de φ convient.□

Remarque. Plaçons-nous dans le cas réel. Pour $\varphi : E \to E$, on a défini une transposée ${}^t\varphi : E^* \to E^*$. On constate que l'isomorphisme $I : E \to E^*$ induit par le produit scalaire transporte la transposée (au sens du produit scalaire) sur l'adjoint (au sens de la dualité). On notera donc souvent ${}^t\varphi$ au lieu de φ^*.

2° Le "théorème spectral"

Définitions : Un endomorphisme est auto-adjoint s'il est égal à son adjoint. Une matrice complexe est hermitienne si elle est égale à son adjointe. Une matrice est symétrique si elle est égale à sa transposée.

Remarque. Un endomorphisme est auto-adjoint, si et seulement si sa matrice dans une base orthonormée est hermitienne/symétrique, si et seulement si sa matrice dans toute base orthonormée est hermitienne/symétrique. (Vérifier l'équivalence !)

Théorème (i) *Un endomorphisme auto-adjoint possède une base orthonormée de vecteurs propres et ses valeurs propres sont réelles.*
(ii) *Une matrice hermitienne (resp. symétrique réelle) possède une base orthonormée de vecteurs propres et ses valeurs propres sont réelles.*

Dans (ii), "orthonormée" se réfère au produit scalaire canonique. Dans le cas réel, les vecteurs propres sont à coordonnées réelles.
Ce théorème est vraiment très utile, même si ça n'apparaît guère dans la suite de ce cours ; il donne quand même, par exemple, l'existence d'axes orthogonaux pour une conique.
Les assertions (i) et (ii) sont clairement équivalentes (prendre une base orthonormée !). On va démontrer (i) par récurrence sur la dimension de E. Pour commencer, on énonce un lemme trivial :

Lemme *Si $\varphi \in \mathcal{L}(E)$ est auto-adjoint, et si $v \in E$ est un vecteur propre de φ, alors $H = \mathrm{Vect}(v)^\perp$ est stable par φ et la restriction de φ à H est encore auto-adjoint.*□

Démonstration du théorème. En dimension 1, il n'y a rien à démontrer. On commence par le cas complexe. Soit v un vecteur propre de φ associé à une valeur propre λ. Notons que λ est réelle, puisque

$$\lambda\langle v, v\rangle = \langle v, \varphi(v)\rangle = \langle \varphi(v), v\rangle = \overline{l}\langle v, v\rangle \quad \text{et} \quad \langle v, v\rangle \neq 0.$$

On vérifie que $\mathrm{Vect}(v)^\perp$ est stable par φ, on y applique l'hypothèse de récurrence et on conclut.
Attaquons le cas réel. La matrice de φ dans une base orthonormée est symétrique, donc hermitienne, donc ses valeurs propres complexes sont en fait réelles. Il en résulte que φ possède un vecteur propre, et la preuve par récurrence marche comme dans le cas complexe.□

Attention ! *Il existe des matrices symétriques complexes non diagonalisables.*

3° Endomorphismes unitaires et isométries

a) Pour $A \in \mathcal{M}_n(\mathbb{C})$, on dit que A est unitaire si $A^* A = \mathrm{Id}$. Alors les valeurs propres de A, donc son déterminant, sont de module 1.
Pour $A \in \mathcal{M}_n(\mathbb{R})$, on dit que A est orthogonale si ${}^tA\, A = \mathrm{Id}$. On a encore : $\det A = \pm 1$.
Notons que A est inversible. On se convainc facilement qu'une matrice est unitaire ou orthogonale si et seulement si c'est la matrice de passage entre deux bases orthonormées.

b) Définition.

Lemme *Soit $\varphi \in \mathcal{L}(E)$. Les conditions suivantes sont équivalentes :*

- $\forall v, w \in E,\ \langle \varphi(v), \varphi(w) \rangle = \langle v, w \rangle$;
- $\forall v \in E,\ ||\varphi(v)|| = ||v||$;
- *φ est l'inverse de son adjoint ;*
- *φ envoie une base orthonormée sur une base orthonormée ;*
- *la matrice de φ dans une base orthonormée est unitaire/orthogonale.*

Définition *Si les conditions du lemme sont réalisées, on dit que φ est une isométrie.*

c) Réduction des isométries réelles (pour mémoire)

Proposition *Un endomorphisme unitaire est diagonalisable dans une base orthonormée.*□

Théorème *Une isométrie d'un espace euclidien possède, dans une base orthonormée convenable, une matrice diagonale par blocs où les blocs sont de la forme*

$$(1), \qquad (-1), \qquad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{avec } \theta \in \mathbb{R}.$$

Démonstration. Par récurrence sur la dimension. C'est trivial en dimension 1. Supposons le théorème vrai dans tout espace de dimension $\leq n-1$. Soit φ une isométrie d'un espace de dimension n, A sa matrice dans une base orthonormée $\mathcal{B}$.
On fixe une valeur propre λ de A. Si elle est réelle, on note P une droite propre. Si elle est complexe, on l'écrit $\lambda = e^{i\theta}$ et on appelle $X \in \mathbb{C}^n$ un vecteur propre de A associé à λ. On note V et W ses parties réelle et imaginaire : ce sont des vecteurs de $\mathbb{R}^n$. On appelle alors v et w les vecteurs qui ont pour coordonnées V et W dans $\mathcal{B}$. On montre aisément que V et W sont linéairement indépendants et on note $P = \mathrm{Vect}(v, w)$.
Il est facile de voir que P et $P^\perp$ sont stables par φ (travailler matriciellement), et que la restriction de φ à P et $P^\perp$ est encore une isométrie, puis on conclut par récurrence.□

4° Endomorphismes normaux (pour mémoire)

Les preuves des deux théorèmes de ce paragraphe se ressemblent énormément. En voici une espèce de généralisation. On dit qu'un endomorphisme φ est normal s'il commute avec son adjoint : $\varphi\,\varphi^* = \varphi^*\,\varphi$. La clé des deux preuves est le lemme suivant :

Lemme *Soit F un sous-espace de E, $\varphi \in \mathcal{L}(E)$ un endomorphisme normal. Si F est stable par φ, alors $F^\perp$ est stable par φ aussi.*□

On en tire avec un raisonnement analogue aux précédents :

Proposition *Tout endomorphisme normal possède une base orthonormée de vecteurs propres.*

Chapitre 6
Espaces euclidiens de dimension 2 et 3

I Isométries et angles en dimension 2

On fixe un plan euclidien P. On note $\mathrm{O}(P)$ le groupe des isométries de P, $\mathrm{SO}(P)$ le groupe des isométries directes (c'est le noyau du déterminant $\det : \mathrm{O}(P) \to \{-1, 1\}$). On note par ailleurs SO_2 le groupe des matrices orthogonales 2×2 de déterminant 1.

1° Rotations du plan, leurs matrices

a) Par définition, une rotation est une isométrie de P de déterminant 1.
b) Le choix d'une base orthonormée $\mathcal{B}$ de P définit un isomorphisme

$$\mathrm{Mat}_{\mathcal{B}} : \mathrm{SO}(P) \xrightarrow{\sim} \mathrm{SO}_2 = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} ; \quad a, b \in \mathbb{R},\ a^2 + b^2 = 1 \right\}.$$

(La description de SO_2 est un calcul simple.)
c) On en déduit facilement les faits cruciaux suivants :

Corollaire (i) SO_2 *est abélien.*
(ii) *Une isométrie a la même matrice dans deux bases de même orientation.*

2° Angles orientés de vecteurs

a) Etant donnés deux couples de vecteurs non nuls (u, v) et (u', v'), on note $(u, v) \sim (u', v')$ s'il existe deux réels $\lambda, \mu > 0$ et une isométrie $\varphi \in \mathrm{SO}(P)$ tels que

$$u' = \lambda\, \varphi(u), \quad v' = \mu\, \varphi(v).$$

C'est un exercice de montrer que $\sim$ est une relation d'équivalence.

Definition. Un angle (orienté de vecteurs) est une classe d'équivalence pour $\sim$. On notera $\mathcal{A}$ l'ensemble des angles.

3° Identification angles orientés-rotations en dimension 2

a) Notons U le cercle unité, formé des vecteurs de norme 1.

Lemme *Fixons $i \in I$. Pour tout $u \in U$, il existe une unique rotation φ telle que $\varphi(i) = u$.*

Remarque. On verra la similitude avec l'assertion familière : Fixons un point $O \in P$. Pour tout point $M \in P$, il existe une unique translation τ telle que $\tau(O) = M$.
Démonstration. Complétons i en une base orthonormée (i, j) de P. Si φ convient, la première colonne de sa matrice dans (i, j) est la colonne des coordonnées ${}^t(a\ b)$ de v. La deuxième colonne est alors bien déterminée. Inversement, cette matrice répond à la question.□
b) On veut définir deux applications entre $\mathcal{A}$ et $\mathrm{SO}(P)$. Pour cela, deux lemmes.

Lemme *Etant donné un angle α, (u, v) un représentant de α, la rotation qui envoie $u/||u||$ sur $v/||v||$ ne dépend que de α.*

Definition. On notera $\mathrm{rot}(\alpha)$ cette rotation.
Démonstration. Tout repose sur la commutativité de $\mathrm{SO}(P)$. Soit (u_1, v_1) et (u_2, v_2) deux représentants de α. Quitte à remplacer u_1 par $u_1/||u_1||$, etc., on peut supposer u_1, v_1, u_2, v_2 de norme 1. Soit φ la rotation qui envoie u_1 sur v_1. On veut montrer que $\varphi(u_2) = v_2$.
Or, par hypothèse, il existe une rotation ψ qui envoie u_1 sur u_2 et v_1 sur v_2. Alors :

$$\varphi(u_2) = \varphi\psi(u_1) = \psi\varphi(u_1) = \psi(u_2) = v_2.\square$$

Lemme *Etant donné une rotation φ et un vecteur non nul u, l'angle $(\widehat{u,\varphi(u)})$ ne dépend que de φ et pas de u.*

DEFINITION. On notera $\mathrm{ang}(\varphi)$ cet angle.

DÉMONSTRATION. Soit u_1, u_2 deux vecteurs non nuls. On peut les supposer de norme 1. Soit ψ la rotation qui envoie u_1 sur u_2. Pour montrer que $(u_1, \varphi(u_1))$ et $(u_2, \varphi(u_2))$ définissent le même angle, on doit montrer que ψ envoie $\varphi(u_1)$ sur $\varphi(u_2)$, ce qui résulte immédiatement de la commutativité de $\mathrm{SO}(P)$.□

Lemme *Les applications* ang *et* rot *sont des bijections réciproques.*

DÉMONSTRATION. Left to the reader...□

4° Addition des angles orientés

Voici une définition un peu trop abstraite, mais efficace, de la somme des angles. On parle de transport de structure. On retrouve la construction géométrique habituelle ("relation de Chasles", si on veut) un peu plus loin.

DEFINITION. Etant donné deux angles α et α', on définit leur somme par :

$$\alpha + \alpha' = \mathrm{ang}(\mathrm{rot}(\alpha) \circ \mathrm{rot}(\alpha')).$$

Lemme *La somme munit l'ensemble des angles d'une structure de groupe abélien, isomorphe au groupe* $\mathrm{SO}(P)$. *Le neutre est l'angle nul $(\widehat{u,u})$, l'opposé de l'angle $(\widehat{u,v})$ est $(\widehat{v,u})$.*□

On peut donner une description géométrique de la somme :

Lemme ("Relation de Chasles") *Soit α et α' deux angles. On fixe $u \in U$. Soit v le vecteur de norme 1 tel que $(\widehat{u,v}) = \alpha$ et w le vecteur de norme 1 tel que $(\widehat{v,w}) = \alpha'$. Alors $\alpha + \alpha' = (\widehat{u,w})$. En particulier, l'angle $(\widehat{u,w})$ ne dépend pas du choix de u.*

DÉMONSTRATION. Puisque $(\widehat{u,v}) = \alpha$, on a : $\mathrm{rot}(\alpha)(u) = v$. De même, $\mathrm{rot}(\alpha')(v) = w$. D'où $\mathrm{rot}(\alpha')\mathrm{rot}(\alpha)(u) = w$, puis $\mathrm{ang}(\mathrm{rot}(\alpha')\mathrm{rot}(\alpha)) = (\widehat{u,w})$ et enfin : $\alpha + \alpha' = \alpha' + \alpha = (\widehat{u,w})$.□

5° Résumé des épisodes précédents

Pour P un plan vectoriel euclidien, on a mis en évidence des bijections entre les trois objets suivants :

- le groupe des isométries directes de P ;
- le groupe des angles orientés de vecteurs de P ;
- le cercle-unité U.

Les deux groupes sont naturellement isomorphes (i.e. on dispose d'un isomorphisme qui ne dépend d'aucun choix, pas même d'une orientation de P). Pour établir une bijection avec le cercle-unité, il faut choisir un point du cercle. Noter l'analogie :

espace vectoriel E	groupe des isométries directes $\mathrm{SO}(P)$
espace affine $\mathcal{E}$	cercle-unité U
origine O	point-base i
translation t de vecteur v	rotation φ d'angle α
$M = t(O) \iff \overrightarrow{OM} = v$	$u = \varphi(i) \iff (\widehat{i,u}) = \alpha$

La structure commune est celle d'un groupe (E ou $\mathrm{SO}(P)$) agissant simplement transitivement sur un ensemble ($\mathcal{E}$ ou U) : deux points quelconques de l'ensemble s'obtiennent l'un à partir de l'autre par l'action d'un unique élément du groupe. En termes snobs, on parle de torseur.

6° Mesure des angles orientés

Supposons connaître le tableau de variations des fonctions cosinus et sinus. On en déduit :

Lemme *Etant donné* $(a,b) \in \mathbb{R}^2$*, tel que* $a^2+b^2=1$*, il existe un unique* $\theta \in\ ]-\pi,\pi]$ *(ou* $\theta \in \mathbb{R}/2\pi\mathbb{Z}$*) tel que* $a=\cos\theta$ *et* $b=\sin\theta$*.*

Démonstration. Il existe $\theta_0 \in [0,\pi]$ tel que $a=\cos\theta_0$. On a alors : $b^2=\sin^2\theta_0$. Si $b \geq 0$, alors $\theta=\theta_0$ convient. Sinon, $\theta=-\theta_0$ convient. L'unicité est analogue.

L'application

$$\begin{array}{ccc} \mathbb{R}/2\pi\mathbb{Z} & \longrightarrow & \mathrm{SO}_2 \\ \theta & \longmapsto & \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \end{array}$$

est canonique. Le choix d'une orientation définit une bijection entre rotations et SO_2. On a également une bijection (canonique) entre angles et rotations. Altogether...

Corollaire *Le choix d'une orientation définit une bijection entre les angles et* $\mathbb{R}/2\pi\mathbb{Z}$*.*

On a ainsi quatre réalisations du "même" groupe : angles, rotations, SO_2, $\mathbb{R}/2\pi\mathbb{Z}$.

7° Autres sortes d'angles

On définit les angles non orientés de vecteurs comme classes d'équivalence de couples de vecteurs non nuls pour la relation :

$$(u,v) \sim_{\mathrm{NV}} (u',v') \iff \exists\lambda,\mu \in \mathbb{R}^{+*},\ \exists\varphi \in \mathrm{O}(P),\ u'=\lambda\,\varphi(u) \text{ et } v'=\varphi(v).$$

Exercice. Montrer que (u,v) et (v,u) définissent le même angle non orienté de vecteurs (normer u et v et utiliser la réflexion d'axe la médiatrice de u et v).

En déduire que (u,v) et (u',v') définissent le même angle non orienté de vecteurs si et seulement si $(\widehat{u,v})=(\widehat{u',v'})$ ou $(\widehat{u,v})=(\widehat{v',u'})$.

Exercice. Montrer que l'ensemble des angles non orientés de vecteurs s'identifie naturellement à l'ensemble des paires (non ordonnées) de demi-droites à isométrie près.

On définit aussi les angles orientés de droites comme classes d'équivalence de couples de vecteurs non nuls pour la relation :

$$(u,v) \sim_{\mathrm{OD}} (u',v') \iff \exists\lambda,\mu \in \mathbb{R}^{*},\ \exists\varphi \in \mathrm{SO}(P),\ u'=\lambda\,\varphi(u) \text{ et } v'=\varphi(v).$$

On définit enfin les angles non orientés de droites comme classes d'équivalence de couples de vecteurs non nuls pour la relation :

$$(u,v) \sim_{\mathrm{ND}} (u',v') \iff \exists\lambda,\mu \in \mathbb{R}^{*},\ \exists\varphi \in \mathrm{O}(P),\ u'=\lambda\,\varphi(u) \text{ et } v'=\varphi(v).$$

Exercice. Dans chaque cas, définir la ou les lignes trigonométriques idoines et la mesure.

8° Isométries négatives

La matrice d'une isométrie indirecte $\varphi \in \mathrm{O}(P) \setminus \mathrm{SO}(P)$ dans une base orthonormée est de la forme

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix};\quad a,b \in \mathbb{R},\ a^2+b^2=1.$$

C'est une réflexion orthogonale, i.e. une symétrie orthogonale par rapport à une droite. On peut montrer qu'une rotation est la composée de deux réflexions.

II Isométries en dimension 3

On fixe un espace euclidien E de dimension 3. Il est souvent commode d'orienter E, mais pas toujours nécessaire. On note $\mathrm{O}(E)$ le groupe des isométries de E, $\mathrm{SO}(E)$ le sous-groupe des isométries directes.

1° Rotations en dimension 3

Définition *Soit E euclidien orienté de dimension 3. Soit $\omega \in E$ un vecteur non nul et $\theta \in \mathbb{R}$. La rotation autour de ω, d'angle θ est l'application linéaire qui a pour matrice :*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

dans une (toute) base orthonormée directe dont le premier vecteur est $\omega/||\omega||$.
Si E n'est pas orienté, une rotation est une application linéaire qui a une matrice de la forme précédente dans une base orthonormée.

En d'autres termes, φ est l'identité sur la droite $\mathrm{Vect}(\omega)$ et c'est une rotation sur le plan orthogonal H. Cette définition a un sens car une rotation a la même matrice dans toute base orthonormée directe de H (ici, l'orientation sur H est induite par le choix de ω).

Remarque. La rotation autour de ω et d'angle θ s'identifie à la rotation autour de $-\omega$ et d'angle $-\theta$. En revanche, elle diffère (en général) de la rotation d'angle $-\omega$ et θ. (Pour voir ceci, constater que si (ω, v_2, v_3) est orthonormée directe, alors $(-\omega, v_3, v_2)$ l'est aussi, mais pas $(-\omega, v_2, v_3)$.) Ainsi, il faut se donner une droite *orientée* et un angle pour définir une rotation.

2° Classification des isométries en dimension 3

Soit φ une isométrie d'un espace euclidien de dimension 3. Le polynôme caractéristique χ_φ de φ est de degré 3, donc il possède une racine réelle. Rappelons que les racines réelles de χ_φ et $\det\varphi$ valent ± 1.
Si χ_φ a une seule racine réelle, on la note ε. Notons qu'alors les autres racines de χ_φ sont complexes conjuguées, donc leur produit est > 0 donc ε est du signe de $\det\varphi$, donc $\varepsilon = \det\varphi$.
Si χ_φ a 3 racines réelles, on pose $\varepsilon = \det\varphi$: c'est une racine de χ_φ. En effet, le produit des trois racines de χ_φ est $\det\varphi$, mais $(-\varepsilon)^3 = -\varepsilon \neq \det\varphi$, donc les racines de χ_φ ne peuvent pas toutes être $-\varepsilon$.
Dans chaque cas, χ_φ a toujours $\varepsilon = \det(\varphi)$ pour racine. Soit alors e_1 un vecteur propre associé à la valeur propre ε. Complétons e_1 en une base orthonormée directe $\mathcal{B} = (e_1, e_2, e_3)$. Notons $D = \mathrm{Vect}(e_1)$ et $H = D^\perp$. Puisque D est stable par φ, il en est de même de H (vérifier !) ; la restriction de φ à H est de plus une isométrie. La matrice de φ dans $\mathcal{B}$ est de la forme

$$\left(\begin{array}{c|cc} \varepsilon & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{B}} \\ 0 & & \end{array}\right),$$

où B est la matrice de la restriction de φ à H. On voit ainsi que $\det\varphi = \varepsilon \det B$, d'où $\det B = 1$ et B est la matrice d'une rotation. On a deux cas :
- si $\det\varphi = 1$, φ est une rotation autour de e_1 ;
- si $\det\varphi = -1$, φ est la composée d'une rotation autour de e_1 et de la réflexion par rapport à $H = \mathrm{Vect}(e_1)^\perp$.

En tout cas, dans une base orthonormée directe convenable, une isométrie a pour matrice

$$A = \left(\begin{array}{c|cc} \varepsilon & 0 & 0 \\ \hline 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{array}\right),$$

où $\varepsilon = \pm 1 = \det A$.

Attention ! *Dans cette matrice, $\theta \in \mathbb{R}$ n'est déterminé qu'à 2π près et au signe près. En effet, θ dépend de l'orientation dans H. Si une orientation de E a été choisie, le choix de θ entre $\pm\theta$ peut être fixé par le choix d'un vecteur de* $\mathrm{Vect}(e_1)$.

Au bilan, toute isométrie directe de E est une rotation, et que toute isométrie indirecte peut s'écrire comme composée d'une rotation et d'une réflexion, l'axe et le plan de celles-ci étant orthogonaux.

3° **Angles orientés et non orientés en dimension** 3

Mimons la définition de la dimension 2. Un angle orienté (resp. non orienté) de vecteurs est une classe d'équivalences de couples de vecteurs non nuls pour la relation $\sim_{OV}$ (resp. $\sim_{NV}$) :

$$(u,v) \sim_{\mathrm{OV}} (u',v') \iff \exists\lambda,\mu \in \mathbb{R}^{+*},\ \exists\varphi \in \mathrm{SO}(P),\ u' = \lambda\,\varphi(u) \text{ et } v' = \varphi(v).$$

$$(u,v) \sim_{\mathrm{NV}} (u',v') \iff \exists\lambda,\mu \in \mathbb{R}^{+*},\ \exists\varphi \in \mathrm{O}(P),\ u' = \lambda\,\varphi(u) \text{ et } v' = \varphi(v).$$

Le nouveau phénomène en dimension 3 est le lemme facile suivant :

Lemme (i) *Etant donné $u,v \in E$, $||u|| = ||v|| = 1$, la rotation d'axe dirigé par $(u+v)/2$ et d'angle π permute u et v.*
(ii) *Les relations $\sim_{OV}$ et $\sim_{NV}$ coïncident.*

Ainsi, les notions d'angle orienté et d'angle non orienté de vecteurs coïncident en dimension 3. En revanche, on ne peut plus identifier les angles et les rotations. Quelle trigonométrie peut-on faire en dimension 3 ?

4° **Produit vectoriel**

Dans ce paragraphe, on choisit une orientation de E.

a) Première version (en coordonnées). Fixons une base orthonormée directe $\mathcal{B}$. Pour $v, v' \in E$, de coordonnées respectives (x,y,z) et (x',y',z') dans E, on note $v \wedge v'$ le vecteur ayant pour coordonnées

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \wedge \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} yz' - zy' \\ zx' - xz' \\ xy' - yx' \end{pmatrix}.$$

b) Deuxième version (géométrique). Le produit vectoriel a les propriétés suivantes :

- $v \wedge v' \in \mathrm{Vect}(v,v')^{\perp}$;
- $||v \wedge v'|| = |\sin(\widehat{v,v'})|.||v||.||v'||$, où $(\widehat{v,v'})$ est l'angle[3] dans le plan $\mathrm{Vect}(v,v')$;
- la famille $(v, v', v \wedge v')$ est une base directe si v et v' ne sont pas colinéaires (et $v \wedge v' = 0$ si v et v' sont colinéaires).

Ces propriétés suffisent pour caractériser $v \wedge v'$, et donnent une deuxième façon de le définir.

c) Produit mixte. Etant donné trois vecteurs $(v,v',w) \in E^3$ (l'ordre compte) et deux bases orthonormées directes $\mathcal{B}$ et $\mathcal{B}'$, on a : $\det_{\mathcal{B}}(v,v',w) = \det_{\mathcal{B}'}(v,v',w)$.
En effet, notons A (resp. A') la matrice dont les colonnes sont les colonnes des coordonnées de v, v' et w dans $\mathcal{B}$ (resp. $\mathcal{B}'$). Notons $P = P_{\mathcal{B}\mathcal{B}'}$ la matrice de changement de base. Comme les deux bases sont supposées orthonormées directes, le déterminant de P vaut 1. De plus, $A = PA'$ donc les déterminants de A et A' sont égaux. Ceci donne un sens à la

Definition. Le produit mixte de trois vecteurs $(v,v',w) \in E^3$ est le déterminant de ces vecteurs dans n'importe quelle base orthonormée directe. On le note $[v,v',w]$. C'est une application trilinéaire et alternée.

[3]Noter que la valeur absolue du sinus ne dépend d'aucune orientation.

d) Troisième version (algébrique). On fixe toujours une base orthonormée directe $\mathcal{B}$. Pour $v, v' \in E$, on considère la forme linéaire $E \to \mathbb{R}$, $w \mapsto [v, v', w]$. Elle est nécessairement de la forme “produit scalaire avec un vecteur”, i.e. il existe un unique vecteur noté $v \wedge v'$ tel que

$$\forall (v, v', w) \in E^3, \quad [v, v', w] = \langle v \wedge v', w \rangle.$$

On vérifie que c'est le même produit vectoriel qu'avant. Seule cette version se généralise dans un espace euclidien de dimension n : étant donné $n-1$ vecteurs, on en fabrique un nouveau.

e) Quelques intérêts du produit vectoriel.
La norme du produit vectoriel $v \wedge v'$ est la surface du parallélogramme supporté par v et v'. La valeur absolue du produit mixte $[v, v', w]$ est le volume du parallélépipède supporté par v, v', w. L'intersection des plans d'équation $ax + by + cz = 0$ et $a'x + b'y + c'z = 0$ est dirigée par le produit vectoriel de ${}^t(a\ b\ c)$ et ${}^t(a'\ b'\ c')$.

5° Reconnaissance des isométries de $\mathbb{R}^3$ (1)

Problème : comment trouver la nature et les éléments caractéristiques d'une isométrie donnée par sa matrice ?
On commence par calculer son déterminant ε. On cherche alors un vecteur v_1 propre pour la valeur propre ε en résolvant un système (on prendra, au moins tacitement, $e_1 = v_1/||v_1||$).
Reste à déterminer l'angle de la rotation de 2°, quand son axe est orienté par v_1.
Tout d'abord, on constate que : $\operatorname{tr} \varphi = \operatorname{tr} A = \varepsilon + 2\cos\theta$. Ceci permet de calculer $\cos\theta$ facilement, donc θ au signe près (et à 2π près).
Pour déterminer le signe de θ, on choisit un vecteur v_2 orthogonal à v_1, on calcule le produit vectoriel $v_2 \wedge \varphi(v_2)$, puis le produit mixte $p = \langle v_2 \wedge \varphi(v_2), v_1 \rangle$. Le signe de $\sin\theta$ est le signe de p. (Vérifier : calculer dans la base orthogonale $(v_1, v_2, v_1 \wedge v_2)$!)

6° Reconnaissance des rotations de $\mathbb{R}^3$ (2)

a) Pour $\omega \in \mathbb{R}^3$ on considère $\Xi_\omega : E \to E$, $v \mapsto \omega \wedge v$. Si $\omega = (a, b, c)$, la matrice de Ξ_ω dans la base canonique est (par définition de $\wedge$) :

$$A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}.$$

Cette matrice est antisymétrique, et toute matrice antisymétrique est de cette forme. Par définition, on a, pour $v \in \mathbb{R}^3$: $Av = \omega \wedge v$. Notons de plus que Ξ_ω a pour noyau $\mathrm{Vect}(\omega)$ si $\omega \neq 0$ et $\mathbb{R}^3$ si $\omega = 0$.
b) Soit à présent ρ une rotation de $\mathbb{R}^3$, R sa matrice dans la base canonique, et posons $A = R - {}^tR = R - R^{-1}$. On voit que A est antisymétrique, donc A est la matrice de Ξ_ω pour un unique $\omega \in \mathbb{R}^3$.
On remarque que si v est un vecteur fixe par ρ, on a : $Rv = v = R^{-1}v$, d'où : $Av = 0$. Il en résulte que v appartient au noyau de Ξ_ω. Si $\omega \neq 0$, ce noyau est de dimension 1 donc ω est fixé par ρ. Si $\omega = 0$, cela reste trivialement vrai.
c) Soit plus précisément ρ la rotation autour de e_1 et d'angle $\theta \in \mathbb{R}$. Par commodité, on suppose $||e_1|| = 1$. On vérifie avec la définition d'une rotation que si $\langle v, e_1 \rangle = 0$, alors : $\rho(v) = \cos(\theta)\, v + \sin(\theta)\, e_1 \wedge v$. Plus généralement, en appliquant cette formule à $v' = v - \langle v, e_1 \rangle\, e_1$, on trouve que :

$$\forall v \in \mathbb{R}^3, \quad \rho(v) = \cos(\theta)\, v + \sin(\theta)\, e_1 \wedge v + (1 - \cos(\theta)) \langle v, e_1 \rangle\, e_1.$$

Il vient alors, pour $v \in \mathbb{R}^3$ quelconque :

$$\rho(v) - \rho^{-1}(v) = (2\sin(\theta)\, e_1) \wedge v.$$

d) Conclusion : soit R la matrice d'une rotation, ω le vecteur tel que $A_\omega = R - {}^tR$. Si $\omega \neq 0$, alors R est la rotation autour de ω et d'angle θ tel que $2\cos\theta + 1 = \operatorname{tr} \rho$ et $\sin\theta = ||\omega||$.
Que se passe-t-il si $\omega = 0$? Pourquoi ne peut-on pas conclure ainsi ?

Chapitre 7
Formes quadratiques

On fixe un espace vectoriel E de dimension finie sur un corps $\mathbb{K}$, de caractéristique $\neq 2$.

I Formes bilinéaires : généralités

1° Définitions

a) Une *forme bilinéaire symétrique* sur E est une application $f : E \times E \to \mathbb{K}$ telle que :
- pour tout $v \in E$, l'application $E \to \mathbb{K}$, $w \mapsto f(v, w)$ est linéaire ;
- pour tout $w \in E$, l'application $E \to \mathbb{K}$, $v \mapsto f(v, w)$ est linéaire ;
- pour tout $(v, w) \in E \times E$, on a : $f(v, w) = f(w, v)$.

Par exemple, le produit scalaire d'un espace euclidien est une forme bilinéaire symétrique...
b) On définit le *noyau* $\mathrm{Ker}(f)$ et le *cône isotrope* $\mathcal{C}_f$ d'une forme bilinéaire symétrique f par :[4]

$$\mathrm{Ker}\, f = \{v \in E \mid \ \forall w \in E,\ f(v, w) = 0\}, \qquad \mathcal{C}_f = \{v \in E \mid \ f(v, v) = 0\}.$$

On dit qu'un vecteur est *isotrope* (pour f) s'il appartient au cône isotrope.

Remarques. Le noyau de f est un sous-espace vectoriel. Le cône isotrope est stable par produit par un scalaire. Le noyau est inclus dans le cône isotrope (vérifier !).
Cependant, en général, le cône n'est pas stable par addition et l'inclusion $\mathrm{Ker}(f) \subset \mathcal{C}_f$ est stricte. Par exemple, on peut prendre $E = \mathbb{R}^2$, et

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix}\right) = xy' + yx',$$

le noyau est réduit à $\{0\}$, et le cône isotrope est formé des axes $x = 0$ et $y = 0$.
c) La *forme quadratique* associée à une forme bilinéaire f sur E est la fonction

$$q : E \longrightarrow \mathbb{K}, \quad v \longmapsto q(v) = f(v, v).$$

Il y a une correspondance bijective entre formes bilinéaires et formes quadratiques. En effet, étant donné une forme quadratique q, on peut exprimer la forme bilinéaire d'où elle provient en fonction de q :

$$\forall v, w \in E, \qquad f(v, w) = \frac{1}{4}\Big(q(v + w) - q(v - w)\Big).$$

On utilise ici l'hypothèse sur la caractéristique de $\mathbb{K}$.

2° Dual et noyau

La donnée d'une forme bilinéaire symétrique détermine une application de E dans son dual E^*, dont on vérifie qu'elle est *linéaire* :

$$\varphi_f : E \longrightarrow E^*,\ v \longmapsto B(v, ?), \qquad \text{où} \quad B(v, ?) : E \to \mathbb{K},\ w \mapsto B(v, w).$$

On constate que le noyau de φ_f est exactement le noyau de f, ce qui justifie cette dénomination. On dit que f est *non dégénérée* si φ_f est un isomorphisme, ce qui revient à dire que le noyau de f est réduit à $\{0\}$ (pourquoi ?). Répétons que le cône isotrope, lui, peut tout à fait être strictement plus gros.

[4] La notation $\mathrm{Ker}(f)$ est standard, mais pas la notation $\mathcal{C}_f$.

3° Matrice d'une forme bilinéaire symétrique

a) Soit $\mathcal{B} = (e_i)_{i=1,\dots,n}$ une base de E, f une forme bilinéaire symétrique sur E. On appelle *matrice de f dans $\mathcal{B}$* la matrice[5] suivante, dont on constate qu'elle est symétrique :

$$A = \mathrm{Mat}_{\mathcal{B}}(f) = \Big(f(e_i, e_j)\Big)_{i,j=1,\dots,n}.$$

b) Formule. Soit $v \in E$ (resp. $w \in E$), $X \in \mathbb{K}^n$ (resp. $Y \in \mathbb{K}^n$) la colonne des coordonnées de v (resp. w) dans $\mathcal{B}$, alors :

$$f(v, w) = {}^tXAY.$$

En effet, il suffit de développer. Si $X = (x_i)_{i=1,\dots,n}$ et $Y = (y_i)_{i=1,\dots,n}$, on a :

$$f(v, w) = f\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i,j=1}^n f(e_i, e_j)\, x_i y_j = {}^tXAY.$$

Si q désigne la forme quadratique associée à f, on a :

$$q(v) = f(v, v) = {}^tXAX = \sum_{i=1}^n \sum_{j=1}^n f(e_i, e_j)\, x_i x_j = \sum_{i=1}^n f(e_i, e_i)\, x_i^2 + \sum_{1 \le i < j \le n} 2f(e_i, e_j)\, x_i x_j.$$

Il y a un réel danger d'oublier le facteur 2...

c) Procédure inverse (le plus utile en pratique).

Proposition *Soit Q une fonction polynômiale homogène de degré 2 sur $\mathbb{K}^n$, c'est-à-dire qu'il existe des scalaires $(b_{ij})_{i,j=1,\dots,n}$ tels que Q soit définie par :*

$$Q : \mathbb{K}^n \longrightarrow \mathbb{K}, \qquad X = (x_i)_{i=1,\dots,n} \longmapsto Q(X) = \sum_{i=1}^n b_{ii}\, x_i^2 + \sum_{1 \le i < j \le n} b_{ij}\, x_i x_j.$$

Alors, Q est une forme quadratique sur $\mathbb{K}^n$ et les b_{ij} sont uniquement déterminés par Q.

On définit alors la matrice de la forme quadratique par :

$$A = (a_{ij})_{i,j=1,\dots,n}, \qquad \text{où} \quad a_{ij} = \begin{cases} b_{ii} & \text{si } i = j, \\ \frac{1}{2} b_{ij} & \text{si } i \neq j. \end{cases}$$

Attention ! *De nouveau, il y a un fort risque d'oublier le coefficient* $1/2$.

Démonstration. On peut vérifier que la fonction F définie par

$$\begin{array}{rccl} F : & \mathbb{K}^n \times \mathbb{K}^n & \longrightarrow & \mathbb{K}, \\ & (X, Y) & \longmapsto & F(X, Y) = \dfrac{Q(X+Y) - Q(X-Y)}{4} \end{array}$$

est bien une forme bilinéaire. Constatons alors que $b_{ij} = F(E_i, E_j)$ (où (E_i) est la base standard de $\mathbb{K}^n$), ce qui montre que la donnée de F détermine les b_{ij}.

[5] Notion standard, notation pas tout à fait standard.

4° Changement de base, congruence

a) Soit $\mathcal{B} = (e_i)$ et $\mathcal{B} = (e'_i)$ deux bases de E, et $P = P_{\mathcal{B},\mathcal{B}'} = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{Id})$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$: ses colonnes sont les coordonnées des e'_j dans $\mathcal{B}$.
Soit alors v et w deux vecteurs de E, X et Y leurs coordonnées respectives dans la base $\mathcal{B}$, X' et Y' leurs coordonnées dans la base $\mathcal{B}'$. On sait que $X = PX'$, et $Y = PY'$. D'où :

$$f(v,w) = {}^tXAY = {}^t(PX')A(PY) = {}^tX'({}^tPAP)Y'.$$

Ceci étant vrai pour tout v, w, on en déduit par le lemme de 3°c) que :

$$\boxed{A' = {}^tPAP,} \qquad \text{c'est-à-dire} \ : \quad \mathrm{Mat}_{\mathcal{B}'}(f) = {}^tP_{\mathcal{B},\mathcal{B}'}\,\mathrm{Mat}_{\mathcal{B}}(f)P_{\mathcal{B},\mathcal{B}'}.$$

Remarquer que c'est différent[6] de la règle de transformation pour les matrices des *applications linéaires* : si on considère A et A' comme les matrices de la même application linéaire dans des bases différentes, et si P est la matrice de passage correspondante, on a : $A' = P^{-1}AP$.

b) Congruence. On définit une relation d'équivalence sur les formes quadratiques ou bilinéaires en disant que deux formes sont *congruentes* si, dans des bases convenables, elles ont la même matrice.
On se pose alors le problème de *classification* des formes à congruence près. C'est en général un problème difficile, qui dépend lourdement du corps sur lequel on travaille, mais on va voir qu'il est trivial sur $\mathbb{C}$ et facile sur $\mathbb{R}$.

5° Bases orthogonales et algorithme de Gauss

a) Bases orthogonales. Etant donné une forme bilinéaire f sur E, on dit qu'une base $(e_i)_{i=1,\ldots,n}$ de E est *orthogonale* si $f(e_i, e_j) = 0$ lorsque $i \neq j$. Voyons la traduction sur l'expression de q en coordonnées.
Notons $(\ell_i)_{i=1,\ldots,n}$ la base duale de $(e_i)_{i=1,\ldots,n}$: ce sont des formes linéaires, uniquement déterminées, telles que $\ell_i(e_j) = \delta_{ij}$ (le δ de Kronecker). Par définition, la i-ème coordonnée d'un vecteur $v \in X$ dans la base $(e_i)_{i=1,\ldots,n}$ est : $\ell_i(v)$. D'où, pour $v \in E$:

$$q(v) = \sum_{i=1}^{n} q(e_i)\,\ell_i(v)^2 + \sum_{1\leq i<j\leq n} f(e_i,e_j)\,\ell_i(v)\ell_j(v) = \sum_{i=1}^{n} a_{ii}\,\ell_i(v)^2 + \sum_{1\leq i<j\leq n} 2a_{ij}\,\ell_i(v)\ell_j(v).$$

D'après la proposition de 3°c), une base est orthogonale si et seulement si l'expression de q en coordonnées est une somme de carrés des ℓ_i.
b) Recherche de bases orthogonales. Par suite, pour trouver une base orthogonale, il suffit d'écrire q comme une somme de carrés de formes linéaires *indépendantes* $\ell_1, \ldots, \ell_r$, disons

$$q = \sum_{i=1}^{r} \alpha_i \ell_i^2.$$

On complète alors $(\ell_1, \ldots, \ell_r)$ en une base $(\ell_1, \ldots, \ell_n)$ du dual de E. La base $(e_1, \ldots, e_n)$, duale de $\mathcal{B}^*$, est alors orthogonale pour q, car la matrice de q dans cette base est, avec une notation évidente : $\mathrm{diag}(\alpha_1, \ldots, \alpha_r, 0, \ldots, 0)$.
c) L'algorithme de Gauss. Il consiste à écrire une forme quadratique comme somme de carrés de formes linéaires, et permet ainsi de trouver des bases orthogonales. On part donc d'une forme

$$Q(X) = \sum_{i=1}^{n} b_{ii}\,x_i^2 + \sum_{1\leq i<j\leq n} b_{ij}\,x_ix_j.$$

[6] Enfin, sauf si P est orthogonale... Voir plus loin, le paragraphe III.

Si $n = 1$, il n'y a rien à faire. Supposons savoir traiter les formes quadratiques en k variables pour $k \leq n-1$.
Premier cas : tous les b_{ii} ne sont pas nuls. Quitte à renuméroter les variables, on peut supposer que $b_{11} \neq 0$. On peut donc écrire Q sous la forme :

$$Q(X) = b_{11}\, x_1^2 + x_1\, L(x_2, \ldots, x_n) + Q'(x_2, \ldots, x_n),$$

où L est une forme linéaire en $x_2, \ldots, x_n$ et Q' est une forme quadratique en les mêmes variables. On factorise alors :

$$Q(X) = b_{11}\left(x_1 + \frac{1}{2b_{11}}L(x_2, \ldots, x_n)\right)^2 - \frac{1}{4b_{11}}L(x_2, \ldots, x_n)^2 + Q'(x_2, \ldots, x_n).$$

On a écrit notre forme Q comme somme du carré d'une forme linéaire et d'une forme quadratique $Q'' = Q' - \frac{1}{4b_{11}}L^2$ qui ne dépend que de $x_2, \ldots, x_n$.
Deuxième cas : tous les b_{ii} sont nuls. Si tous les b_{ij} sont nuls, il n'y a rien à faire. On suppose donc que l'un d'entre eux n'est pas nul, par exemple $b_{12} \neq 0$. On écrit alors

$$Q(X) = b_{12}\, x_1 x_2 + x_1\, L_2(x_3, \ldots, x_n) + x_2\, L_1(x_3, \ldots, x_n) + Q'(x_3, \ldots, x_n),$$

où L_1 et L_2 sont des formes linéaires, et Q' est une forme quadratique en les variables $x_3, \ldots, x_n$. On factorise alors :

$$Q(X) = b_{12}\left(x_1 + \frac{1}{b_{12}}L_1\right)\left(x_2 + \frac{1}{b_{12}}L_2\right) - \frac{1}{b_{12}^2}L_1L_2 + Q',$$

ce qu'on peut écrire, en vertu de : $uv = \frac{1}{4}(u+v)^2 - \frac{1}{4}(u-v)^2$:

$$Q(X) = \frac{b_{12}}{4}\left(x_1 + x_2 + \frac{1}{b_{12}}L_1 + \frac{1}{b_{12}}L_2\right)^2 - \frac{b_{12}}{4}\left(x_1 - x_2 + \frac{1}{b_{12}}L_1 - \frac{1}{b_{12}}L_2\right)^2 - \frac{1}{b_{12}^2}L_1L_2 + Q',$$

et on a exprimé Q comme somme des carrés de deux formes linéaires indépendantes et indépendantes de $x_3, \ldots, x_n$, et d'une forme quadratique qui ne dépend que de $x_3, \ldots, x_n$.

Dans les deux cas, on a réduit le nombre de variables, ce qui permet d'embrayer la récurrence.
d) Traduction matricielle. On a montré :

Proposition *Quel que soit le corps $\mathbb{K}$, pour toute matrice symétrique A à coefficients dans $\mathbb{K}$, il existe une matrice inversible P triangulaire ou triangulaire par blocs 2×2 telle que ${}^tPAP$ est diagonale.*

Ce résultat est, pour les matrices symétriques, une sorte d'analogue du "théorème du rang" (version matricielle, théorème 3° du premier chapitre) : pour toute matrice rectangulaire A, il existe P et Q carrées inversibles telles que $Q^{-1}AP$ soit [...]
D'ailleurs, on voit bien que le principe de preuve est très comparable, et porte le même nom : algorithme de Gauss.

Attention ! *Les coefficients qui apparaissent sur la diagonale n'ont* aucune raison *d'être les valeurs propres de la matrice de la forme quadratique. Voir par exemple les paragraphes II et III pour s'en convaincre.*
En particulier, ce théorème ne dit pas *que toute matrice symétrique est diagonalisable, car ${}^tPAP$ ne saurait être confondu avec $P^{-1}AP$ en général.*
Par exemple, la matrice $\begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$ *n'est pas diagonalisable.*

II Classification des formes bilinéaires sur $\mathbb{C}$

La situation est extrêmement simple, car tout complexe est un carré. On vient de montrer avec l'algorithme de Gauss que toute forme quadratique q sur E s'écrit, dans une certaine base $(e'_1,\ldots,e'_n)$ et pour $r \geq n$ et $\alpha_1,\ldots,\alpha_r \in \mathbb{C}\setminus\{0\}$ convenables :

$$q(v) = \sum_{i=1}^{r} \alpha_i\, {x'_i}^2 \qquad \text{si} \quad v = \sum_{i=1}^{n} x'_i\, e'_i.$$

Pour $i = 1,\ldots,r$, choisissons une racine carré β_i de α_i, et pour $i \geq r+1$, posons $\beta_i = 1$. Posons alors $e_i = \beta_i\, e'_i$. La i-ème coordonnée de v dans $(e_i)_{i=1,\ldots,n}$ est : $x_i = \frac{1}{\beta_i}\, x'_i$. Par suite :

$$q(v) = \sum_{i=1}^{r} {x_i}^2 \qquad \text{si} \quad v = \sum_{i=1}^{n} x_i\, e_i.$$

Ainsi, à congruence près, une forme quadratique est classée par son rang. Pas palpitant.

III Formes bilinéaires sur $\mathbb{R}$

Sur $\mathbb{R}$, la situation est plus subtile, car il y a deux types de réels non nuls modulo les carrés.

1° Théorème d'inertie de Sylvester

Théorème *Soit E de dimension finie sur $\mathbb{R}$ et q une forme quadratique sur E.*
(i) *Il existe une base $(e_i)_{i=1,\ldots,n}$ de E, deux entiers $r, s \in \mathbb{N}$ tels que*

$$q(v) = \sum_{i=1}^{r} x_i^2 - \sum_{j=r+1}^{r+s} x_j^2 \qquad \textit{si} \quad v = \sum_{i=1}^{n} x_i\, e_i \in E.$$

(ii) *Les entiers r et s ne dépendent que de la forme quadratique, et pas de la base orthogonale dans laquelle une telle écriture a lieu.*

Remarque. Il faut comprendre ce théorème comme un résultat de *classification* des formes quadratiques sur $\mathbb{R}$, par la *signature* (r, s) de la forme : c'est, par définition, le couple formé par le nombre de “carrés positifs” et le nombre de “carrés négatifs”.

Démonstration. L'assertion (i) est une conséquence facile de l'algorithme de Gauss. Soit $(e'_i)_{i=1,\ldots,n}$ une base orthogonale pour q. Quitte à permuter les e'_i, on peut supposer que

$$\alpha_i = q(e_i) > 0 \text{ si } i \leq r, \quad \alpha_i = q(e_i) < 0 \text{ si } r+1 \leq i \leq r+s, \quad q(e_i) = 0 \text{ si } i > r+s,$$

ce qui définit deux naturels r et s. On pose alors :

$$e_i = \sqrt{\alpha_i}\, e'_i \text{ si } i \leq r, \quad e_i = \sqrt{-\alpha_i}\, e'_i \text{ si } r+1 \leq i \leq r+s, \quad e_i = e'_i \text{ si } i > r+s,$$

et, comme dans II, la base ainsi construite convient.

La preuve de (ii) est magnifique. Il s'agit de caractériser les entiers r et s qui apparaissent dans (i) de façon intrinsèque, i.e. sans utiliser la base (e_i). On va montrer que :

$$\begin{cases} r = \max\{\dim F,\ F \text{ sous-espace de } E \text{ sur lequel } q \text{ est définie positive}\}, \\ s = \max\{\dim F,\ F \text{ sous-espace de } E \text{ sur lequel } q \text{ est définie négative}\}. \end{cases}$$

“Rappelons” que q est définie négative sur F si $q(v) \leq 0$ pour tout $v \in F$, avec égalité seulement si $v = 0$. On donne la preuve uniquement pour s, l'autre est inutile (remplacer q par $-q$). Notons S la dimension maximale d'un sous-espace sur lequel q est définie négative.

Tout d'abord, vu que q est définie négative sur le sous-espace $F = \mathrm{Vect}(e_{r+1}, \ldots, e_{r+s})$, et que de dimension s, on a : $s \leq S$ (maximalité de S).
Par ailleurs, soit G est un sous-espace de E de dimension $d \geq s+1$. Notons $H = \mathrm{Vect}(e_1, \ldots, e_r, e_{r+s+1}, \ldots, e_n)$. On a : $\dim G + \dim H > \dim E$, donc l'intersection $G \cap H$ est de dimension ≥ 1. Soit $v \in G \cap H$, $v \neq 0$. Comme $v \in H$, v s'écrit

$$0 \neq v = x_1 e_1 + \cdots + x_r\, v_r + x_{r+s+1}\, e_{r+s+1} + \cdots + x_n\, e_n, \quad \text{d'où} \quad q(v) = \sum_{i=1}^{r} x_i^2 \geq 0.$$

Ceci montre q ne peut pas être définie négative sur G, donc que $S < s+1$. Au bilan : $s = S$.

2° Diagonalisation dans une base orthonormée

a) Ici, on munit E d'un produit scalaire euclidien et d'une forme quadratique q. Fixons une base orthonormée $(e_i)_{i=1,\ldots,n}$ de E. On a déjà remarqué que la matrice de q dans cette base est symétrique. Or, une matrice symétrique réelle[7] est diagonalisable dans une base orthonormée $(e'_i)_{i=1,\ldots,n}$. Il existe donc une matrice orthogonale P telle que $D = P^{-1}AP$ soit diagonale. Or, dire que P est orthogonale, c'est dire : $P^{-1} = {}^tP$. Ainsi, D est la matrice de la forme quadratique q dans la base $(e'_i)_{i=1,\ldots,n}$, qui est donc orthogonale pour q. On a prouvé :

Théorème *Pour toute forme quadratique sur un espace euclidien de dimension finie, il existe une base qui est à la fois orthonormée pour le produit scalaire et pour la forme quadratique.*

Comme application, notons que c'est de ce théorème que découle l'existence des axes d'une conique ou d'une quadrique, par exemple.

b) Remarque. Algorithmiquement, il est beaucoup plus difficile de mettre en œuvre ce théorème que la réduction de Gauss, mais l'information obtenue est beaucoup plus précise. Ici, on doit trouver/obtient les valeurs propres de la matrice ; en revanche, par la réduction de Gauss, on n'obtient que *les signes* des valeurs propres. En particulier, si on s'intéresse seulement au *genre* d'une conique, c'est-à-dire à la signature de la forme quadratique associée, pas besoin de diagonaliser la matrice de la forme quadratique : la réduction de Gauss suffit.

[7]Insistons : ceci *ne marche pas* sur les complexes !

Chapitre 8
Coniques

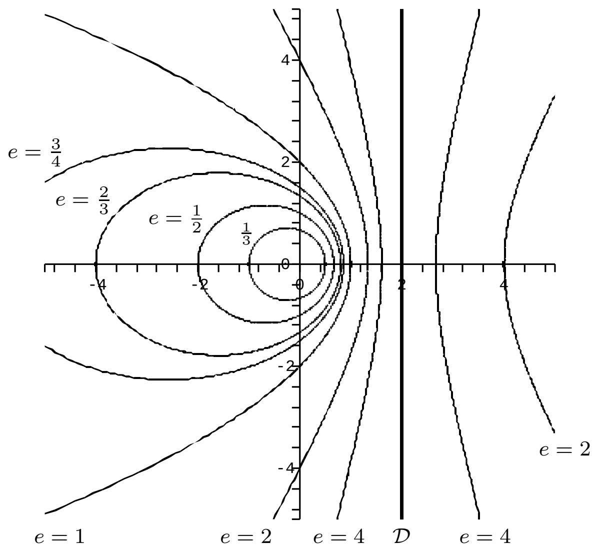
I Foyer et directrice

1° Définition d'une conique par foyer et directrice

On fixe un plan affine euclidien orienté $\mathcal{P}$, un point F, une droite $\mathcal{D}$ tels que $F \notin \mathcal{D}$, et un réel $e > 0$. On associe à ces données la conique

$$\mathcal{C} = \{M \in \mathcal{P},\ MF = e\, d(M, \mathcal{D})\},$$

où $d(M, \mathcal{D})$ désigne la distance de M à la droite $\mathcal{D}$. Si H est le projeté orthogonal de M sur $\mathcal{D}$, on a bien sûr : $d(M, \mathcal{D}) = MH$.
On appelle K le projeté orthogonal de F sur $\mathcal{D}$, on pose $i = \overrightarrow{FK}/FK$ et j le vecteur tel que (F, i, j) soit un repère orthonormé direct. Les coordonnées seront relatives à ce repère.
Voici, lorsque F et $\mathcal{D}$ sont fixés, l'allure des coniques lorsque e varie.



2° Coordonnées polaires

Il y a une ambiguïté quand on parle de coordonnées polaires.
a) Le sens standard est le suivant. Soit M un point de $\mathcal{P}$, (x, y) ses coordonnées dans (F, i, j). Il existe un réel *positif* $\rho \geq 0$ unique et un réel θ, unique à 2π près si $M \neq F$, tels que l'on ait :

$$(*) \qquad \overrightarrow{FM} = \rho \cos\theta\, i + \rho \sin\theta\, j \quad \text{ou (c'est pareil)} \quad \begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta. \end{cases}$$

Le couple (ρ, θ) forme, par définition, les coordonnées polaires de M dans le repère (F, i, j). Géométriquement, on trouve les coordonnées polaires de M ainsi : $\rho = FM$ et θ est une mesure de l'angle orienté de vecteurs $(i, \overrightarrow{FM})$.

b) Cependant, on peut avoir intérêt à autoriser ρ à être négatif. Dans ces conditions, on peut appeler "quasi-coordonnées polaires"[8] de $M \neq F$ tout couple $(\rho, \theta) \in \mathbb{R}^* \times (\mathbb{R}/2\pi\mathbb{Z})$

[8] Monier appelle un tel couple "système de coordonnées polaires de M".

tel que la relation $(*)$ ci-dessus soit satisfaite. Dans ces conditions, on perd l'unicité des coordonnées polaires. En effet, les points ayant pour quasi-coordonnées polaires (ρ,θ) et $(-\rho,\theta+\pi)$ coïncident :

$$\rho\cos\theta\, i+\rho\sin\theta\, j=(-\rho)\cos(\theta+\pi)\, i+(-\rho)\sin(\theta+\pi)\, j.$$

Géométriquement, θ change d'interprétation et ρ devient une mesure algébrique :

$$\theta=(\widehat{i,u_\rho})\quad\text{et}\quad\rho=\langle\overrightarrow{FM},u_\rho\rangle,\qquad\text{où } u_\rho=\cos\theta\, i+\sin\theta\, j.$$

c) A présent, soit $f:\mathbb{R}\to\mathbb{R}$ une fonction. On appelle courbe d'équation polaire

$$\rho=f(\theta)$$

dans le repère (F,i,j), l'ensemble des points $M(\theta)$ pour $\theta\in\mathbb{R}$, où $M(\theta)$ est défini par :

$$\overrightarrow{FM(\theta)}=f(\theta)\cos\theta\, i+\sin\theta\, j.$$

On constate que si on impose à f d'être positive, $(f(\theta),\theta)$ sont les coordonnées polaires de $M(\theta)$ (en tout point tel que $f(\theta)\neq 0$), et qu'en général, $(f(\theta),\theta)$ sont seulement des quasi-coordonnées polaires de M.

3° Equation polaire des coniques

a) On veut montrer que la conique $\mathcal{C}$ de 1° peut être décrite par l'équation polaire suivante[9] :

$$\rho=\frac{p}{1+e\cos\theta},\qquad\text{où } p=e.d(F,\mathcal{D})=e\,FK.$$

Soit M un point, (x,y) ses coordonnées cartésiennes, (ρ,θ) des quasi-coordonnées polaires. On commence par calculer quelques coordonnées dans (F,i,j) :

$$F:\begin{pmatrix}0\\0\end{pmatrix},\quad\mathcal{D}:x=c,\quad M:\begin{pmatrix}x\\y\end{pmatrix},\quad H:\begin{pmatrix}c\\y\end{pmatrix},\quad\text{où } c=d(F,\mathcal{D})=FK.$$

Ainsi, on a :

$$\begin{aligned}M\in\mathcal{C}\iff\ & MF=e\,MH\iff\sqrt{x^2+y^2}=e|x-c|\\ \iff\ & |\rho|=e|\rho\cos\theta-c|\iff\left\{\begin{array}{l}(1+e\cos\theta)\rho=ec\\ \text{ou}\\ (1-e\cos\theta)\rho=-ec\end{array}\right.\end{aligned}$$

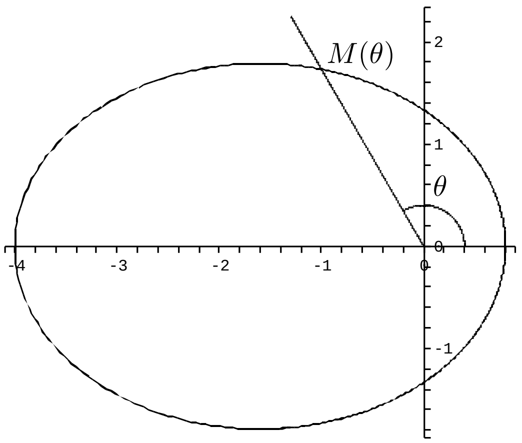
Or, compte tenu du fait que les points de quasi-coordonnées polaires (ρ,θ) et $(-\rho,\theta+\pi)$ coïncident, les courbes d'équations polaires

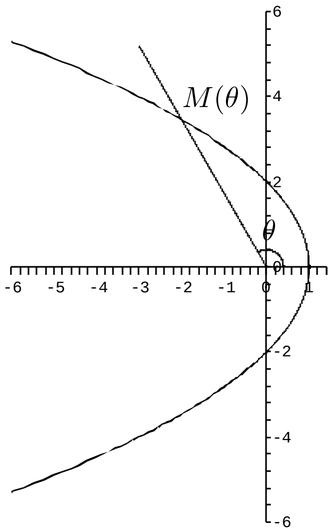
$$\rho=\frac{-ec}{1-e\cos\theta}\quad\text{et}\quad\rho=\frac{ec}{1+e\cos\theta}$$

coïncident ! Si on pose $f(\theta)=ec/(1+e\cos\theta)$ et $g(\theta)=-ec/(1-e\cos\theta)$, le point de paramètre θ de la courbe $\rho=f(\theta)$ coïncide avec le point de paramètre $\theta+\pi$ de la courbe $\rho=g(\theta)$.
Par commodité, on ne garde que l'équation qui contient les signes $+$. On a obtenu l'équation polaire souhaitée.

[9] Attention à ne pas faire vos ρ comme vos p !

b) Pour $e \leq 1$, on a : $1 + e\cos\theta > 0$ (sauf si $e = 1$ et $\theta = \pi\ [2\pi]$), donc on est bien en coordonnées polaires classiques ($\rho > 0$). Toute demi-droite d'origine F coupe $\mathcal{C}$ en un point et un seul (sauf $F + \mathbb{R}^- i$ lorsque $e = 1$).

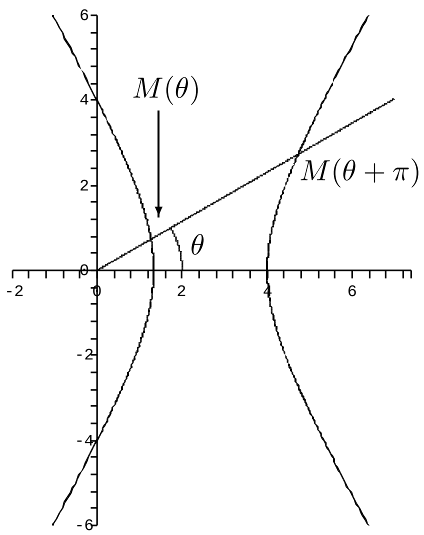




En revanche, pour $e > 1$, l'intérêt d'autoriser des $\rho < 0$ apparaît. En effet, parmi les demi-droites contenant F et dirigées par $u_\rho = \cos\theta\, i + \sin\theta\, j$:

- certaines ne contiennent aucun point de $\mathcal{C}$;
- certaines en contiennent deux !

En revanche, chaque droite contenant F contient exactement deux points de $\mathcal{C}$ (à deux exceptions près, correspondant aux points où $1+e\cos\theta = 0$). L'intérêt de l'équation polaire proposée est de paramétrer la conique de façon uniforme par rapport à θ, au très faible prix de passer des "coordonnées polaires" aux "quasi-coordonnées polaires".



Ainsi, l'équation des coniques en polaires est complètement uniforme, alors que leur allure est assez variable.

Chapitre 9
Google et le théorème de Perron

On donne ici une version "académique" de l'algorithme `PageRank` utilisé par Google pour classer les pages web "par ordre d'importance", de manière à les présenter dans l'ordre le plus pertinent à l'issue d'une recherche.
L'idée est de se promener au hasard sur le web, en suivant des liens. On classe alors les pages selon la probabilité de s'y trouver. Il s'agit donc de donner un sens à l'expression "probabilité d'être à une page donnée" et un moyen de la calculer.
Les méthodes utilisées seront le vocabulaire de la théorie des graphes, un zeste de probabilités et un théorème d'algèbre linéaire attribué à Perron et Frobenius. L'idée probabiliste sous-jacente est celle de *chaîne de Markov (finie)*, mais on l'utilisera tacitement.

1° Des graphes

a) Un *graphe orienté (fini)* est la donnée
- d'un ensemble fini I dont les éléments sont appelés *sommets* ;
- d'un ensemble fini F dont les éléments sont appelés *flèches* ;
- de deux application $s, b : F \to I$, appelées respectivement *source* et *but*.

b) La donnée d'une partie $A \subset I \times I$ détermine F, s et b par la recette suivante : on prend $F = A$, et pour $\alpha = (i,j) \in F$, on pose $s(\alpha) = i$ et $b(\alpha) = j$.
L'intérêt de la présentation du a) est d'autoriser plusieurs flèches entre deux sommets fixés. Les puristes pourront s'en offusquer.

c) On modélise alors le web par un graphe :
- on prend pour sommets l'ensemble I des pages du web ;
- on prend pour flèches l'ensemble F des liens hypertextes entre deux pages ;
- étant donné un lien $\alpha \in F$, sa source $s(\alpha)$ est la page où se trouve le lien, et son but $b(\alpha)$ est la page vers laquelle pointe le lien.

d) Quitte à numéroter les sommets, on supposera désormais que $I = \{1, \ldots, n\}$ pour $n \in \mathbb{N}^*$ convenable.
La *matrice d'adjacence du graphe* est la matrice $A = (a_{ij})_{i,j=1,\ldots,n}$ définie par :

$$\forall i, j \in \{1, \ldots, n\}, \qquad a_{ij} = \operatorname{card}\{\alpha \in F,\ s(\alpha) = j \text{ et } b(\alpha) = i\}.$$

En mots, a_{ij} est le nombre de flèches de j vers i. (Attention à l'ordre des indices.)

Remarque. A permutation/renumérotation des arêtes et des flèches près, la matrice d'adjacence détermine le graphe.

e) Matrice d'adjacence et chemins
Pour $i, j \in I$, on définit un chemin de longueur $\ell \geq 1$ de j vers i comme une suite de flèches $(\alpha_1, \ldots, \alpha_\ell) \in F^\ell$ telles que $s(\alpha_2) = b(\alpha_3), \ldots,\ s(\alpha_\ell) = b(\alpha_{\ell-1})$.

Lemme *Pour $i, j \in I$, le nombre de chemins de longueur $\ell \geq 1$ de j vers i est le coefficient d'indice (i,j) de A^ℓ, où A est la matrice d'adjacence.*

Démonstration. C'est évident si $\ell = 1$ par définition de A. Supposons que ce soit vrai pour un certain $\ell \geq 1$. Notons $A^\ell = (b_{ij})_{i,j}$. Un chemin de j vers i de longueur $\ell + 1$ est une suite $(\alpha_1, \ldots, \alpha_{\ell+1}) \in F^{\ell+1}$ telle que $(\alpha_1, \ldots, \alpha_\ell)$ soit un chemin de longueur ℓ partant de j (de source $s(\alpha_1) = j$), et $\alpha_{\ell+1} \in F$ soit une flèche arrivant en i (de but $b(\alpha_{\ell+1}) = i$). Le nombre de ces chemins est donc :

$$\sum_{k \in I} \operatorname{card}\left\{ \begin{array}{c} \text{chemins de longueur } \ell \\ \text{de } j \text{ vers } k \end{array} \right\} \times \operatorname{card}\left\{ \begin{array}{c} \text{flèches} \\ \text{de } k \text{ vers } i \end{array} \right\} = \sum_{k=1}^{n} b_{kj} a_{ik} :$$

on reconnaît là le coefficient d'indice (i,j) de $A.B = A^{\ell+1}$.□

Corollaire *Il existe un chemin de j vers i si et seulement si, pour $\ell \in \mathbb{N}$ assez grand, le coefficient d'indice (i,j) de $\mathrm{Id} + A + \cdots + A^\ell$ est strictement positif.*

2° Marche au hasard sur un graphe

a) Idée

On se promène au hasard sur le graphe avec la règle suivante : si on est à la page $j \in I$, on choisit une page i vers laquelle il y a un lien, avec une probabilité proportionnelle au nombre de liens de la page j vers la page i. Si jamais il n'y a pas de lien issu de la page j, on décide de choisir une page au hasard avec probabilité uniforme.
Ainsi, la probabilité d'atterrir à la page i, sachant qu'on est à la page j est :

$$s_{ij} = \begin{cases} \dfrac{a_{ij}}{\sum_{k=1}^n a_{kj}} & \text{si } \sum_{k=1}^n a_{kj} \neq 0, \\ \dfrac{1}{n} & \text{sinon.} \end{cases}$$

On associe au graphe la matrice suivante, qui donne les *transitions élémentaires* d'une page à l'autre :

$$S = (s_{ij})_{i,j=1,\ldots,n}.$$

Remarque : La matrice S est *stochastique* (selon les colonnes), au sens suivant (vérifier !) :

$$\forall j \in \{1,\ldots,n\}, \qquad \sum_{i=1}^n s_{ij} = 1.$$

b) Marche au hasard : formalisation

Soit $(\Omega, \mathcal{B}, P)$ un univers probabilisé. On veut modéliser une marche au hasard dans laquelle la position au temps $N+1$ ne dépend que de la position au temps N, et est déterminée par les considérations précédentes. On introduit donc une suite de variables aléatoire $(\Pi_N)_{N\in\mathbb{N}}$ (Π pour "page"), à valeurs dans $\{1,\ldots,n\}$, et on fait l'hypothèse suivante :

$$\forall i,j \in \{1,\ldots,n\},\ \forall N \in \mathbb{N}, \qquad P(\Pi_{N+1} = i | \Pi_N = j) = s_{ij} = \frac{a_{ij}}{\sum_{k=1}^n a_{kj}}.$$

Le terme savant pour une telle suite de variables aléatoires est : *chaîne de Markov finie.*

La "formule des probabilités totales" permet d'écrire :

$$(*) \qquad \forall N \in \mathbb{N},\ \forall i \in \{1,\ldots,n\}, \qquad P(\Pi_N = i) = \sum_{j=1}^n P(\Pi_{N+1} = i | \Pi_N = j)\, P(\Pi_N = j).$$

Pour tout $N \in \mathbb{N}$, on définit un vecteur $X^{(N)} = (x_i^{(N)})_{i=1,\ldots,n}$ par :

$$\forall i = 1,\ldots,n, \qquad x_i^{(N)} = P(X_N = i).$$

L'égalité $(*)$ s'écrit :

$$\forall N \in \mathbb{N}, \qquad X^{(N+1)} = SX^{(N)},$$

d'où l'on tire :

$$\forall N \in \mathbb{N}, \qquad X^{(N)} = S^N X^{(0)}.$$

But : Evaluer (et classer par ordre décroissant) les $x_i^{(N)}$ pour N grand.

c) Modification du processus (magouille !)
Vous avez compris ce qui précède ? Alors, on change de cadre. On constate qu'en fait, quand un surfeur se promène sur le net, il lui arrive de passer d'une page à une autre en entrant directement l'adresse de cette dernière, sans suivre un lien. On modélise ce comportement de la façon suivante.
On introduit la matrice J dont les coefficients sont égaux à $1/n$. Ce serait la matrice S du graphe du web s'il y avait sur chaque page, un lien vers toutes les autres pages du web. Se promener au hasard sur un tel graphe, cela revient à choisir à chaque instant N une page au hasard avec probabilité uniforme parmi les n disponibles.
On suppose que 85% du temps, un surfeur suit la marche au hasard du paragraphe précédent, et que 15% du temps, il choisit une page au hasard avec loi uniforme. (Le choix de ces pourcentages est arbitraire.) La formule des probabilités totales légitime l'introduction et l'interprétation de la matrice suivante :

$$T = 0{,}85\, S + 0{,}15\, C.$$

Si $x_i^{(N)}$ est la probabilité d'être à la page i à l'instant N, et si $X^{(N)} = (x_i^{(N)})_{i=1,\ldots,n}$, on a donc :

$$\forall N \in \mathbb{N}, \qquad X^{(N+1)} = TX^{(N)},$$

ou encore :

$$\forall N \in \mathbb{N}, \qquad X^{(N)} = T^N X^{(0)}.$$

But : Evaluer (et classer par ordre décroissant) les $x_i^{(N)}$ pour N grand.

3° L'algèbre linéaire à l'aide : théorème de Perron

a) On "rappelle" que le rayon spectral d'une matrice T à coefficient complexe est le plus grand module d'une valeur propre de T :

$$\rho(T) = \max\left\{|\lambda|,\ \lambda \in \mathbb{C} \text{ et } \det(T - \lambda \mathrm{Id}) = 0\right\}.$$

Théorème (Perron) *Soit $T \in \mathcal{M}_n(\mathbb{R}^{+*})$ une matrice carrée dont tous les coefficients sont strictement positifs. Alors :*
(i) $\rho(T)$ est une valeur propre simple de T ;
(ii) toute autre valeur propre complexe $\lambda \neq \rho(T)$ satisfait : $|\lambda| < \rho(T)$;
(iii) il existe un unique vecteur propre de valeur propre $\rho(T)$ dont tous les coefficients sont strictement positifs, et donc la somme des coefficients vaut 1.

On appelle $\rho(T)$ la valeur propre de Perron-Frobenius de T, et le vecteur propre de (iii) le vecteur propre de Perron-Frobenius. Il y aura peut-être une preuve un jour dans ces notes.

b) Rayon spectral des matrices stochastiques
On dira que la matrice $T = (t_{ij})_{i,j=1,\ldots,n}$ est strictement positive si tous ses coefficients sont strictement positifs, et qu'elle est stochastique si :

$$\forall i, j = 1, \ldots, n,\ t_{ij} > 0.$$

On notera $||.||$ la "norme 1" sur $\mathbb{R}^n$:

$$\forall X = (x_i)_{i=1,\ldots,n} \in \mathbb{R}^n, \qquad ||X|| = \sum_{i=1}^{n} |x_i|.$$

Lemme *Si T est strictement positive et stochastique, et si X est son vecteur de Perron-Frobenius, alors : $||TX|| \leq ||X||$.*

Démonstration. Facile :

$$\sum_{i=1}^{n}\left|\sum_{j=1}^{n} t_{ij}x_j\right| \leq \sum_{i=1}^{n}\sum_{j=1}^{n} t_{ij}|x_j| \leq \sum_{j=1}^{n}\left(\sum_{i=1}^{n} t_{ij}\right)|x_j| \leq \sum_{j=1}^{n}|x_j| = ||X||.$$

Proposition *Si T est strictement positive et stochastique, alors : $\rho(T) = 1$.*

Démonstration. D'une part, si X est un vecteur propre de T, et λ est la valeur propre associée, on a d'après le lemme :

$$|\lambda|\ ||X|| = ||TX|| \leq ||X||, \quad \text{d'où} \quad |\lambda| \leq 1, \quad \text{puis} \quad \rho(T) \leq 1.$$

D'autre part, le vecteur ${}^t(1\ 1 \cdots 1)$ est un vecteur propre de la transposée de T, associé à la valeur propre 1, donc 1 est valeur propre de T, donc $\rho(T) \geq 1$.□

c) Dynamique d'une chaîne de Markov

Proposition *Soit $T \in \mathcal{M}_n(\mathbb{R})$ une matrice strictement positive et stochastique. Alors, pour tout vecteur $X^{(0)} \in (\mathbb{R}^+)^n$ à coordonnées positives, tel que $||X^{(0)}|| = 1$, la suite $(T^N X^{(0)})_{N\in\mathbb{N}}$ converge vers le vecteur de Perron-Frobenius de T.*

Démonstration. On triche un peu en supposant que T est diagonalisable, ce qui est presque sûrement le cas, et qui n'est qu'une façon de se simplifier la vie de toute façon.
Bref. Soit $\lambda_1, \ldots, \lambda_n$ les valeurs propres de T, numérotées de sorte que $\lambda_1 = 1$. Pour $i \geq 2$, on a donc : $|\lambda_i| < 1$. Il existe une matrice inversible à coefficients complexes P, dont la première colonne est le vecteur de Perron-Frobenius V, telle que

$$T = P\begin{pmatrix} 1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}P^{-1}.$$

Comme $\lambda_2, \ldots, \lambda_n$ sont de module < 1, on a :

$$T^N X^{(0)} \overset{N\to+\infty}{\longrightarrow} P\begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}P^{-1}X^{(0)}.$$

On en déduit sans peine que $T^N X^{(0)}$ tend vers un multiple de la première colonne de P. Comme $T^N X^{(0)}$ est à coefficients strictement positifs, et comme $||TX|| = ||X||$ si $X \in (\mathbb{R}^+)^n$ (vérifier), on en déduit que la limite est exactement le vecteur de Perron-Frobenius.

Remarque : L'introduction de C dans le processus (en 2°c)) sert d'une part à assurer que T soit strictement positive, d'autre part à contrôler (majorer) $|\lambda_2|$ pour assurer que la suite $(T^N X^{(0)})$ converge assez vite.

d) Conclusion

Si le vecteur propre de Perron-Frobenius de T est ${}^t(p_1 \ \cdots \ p_n) \in (\mathbb{R}^{+*})^n$, on a montré que la probabilité d'être à la page i en un instant N assez grand est proche de p_i. On en déduit

- un classement des pages selon la valeur de p_i ;
- une méthode de calcul "en principe" des p_i par le calcul des puissances de T.

Le problème informatique est énorme : n est le nombre de pages du web, assurément plusieurs millions et sans doute plus ! Dans la pratique, il semble que Google fasse travailler 600 à 700 ordinateurs en parallèle. Leur capacité à le faire est (une de) leur(s) grande(s) force(s).

Questionnaire dogmatique

1° Vecteurs

Parmi les objets suivants, lesquels peuvent être des éléments d'un espace vectoriel :

fonction réelle, nombre réel, vecteur, nombre complexe, matrice, blorp ?

2° Vocabulaire

Parmi les mots et expressions suivants, deux (ou trois) nécessitent un *théorème* pour prendre un sens : lesquels ?

- espace vectoriel, sous-espace vectoriel, supplémentaires ;
- application linéaire, image, noyau ;
- combinaison linéaire ;
- famille libre, famille génératrice, base, dimension, rang ;
- déterminant ;
- valeur propre, vecteur propre, espace propre.

3° Classification

Comment classe-t-on les espaces vectoriels ? les applications linéaires ? les endomorphismes ?

4° Applications linéaires

On peut définir une application linéaire par l'image de tous les vecteurs de l'espace de départ. Donner trois autres façons de caractériser une application linéaire.

5° Opérations élémentaires sur les rangées

Donner cinq types de problèmes pour lesquelles des opérations élémentaires sur les rangées d'une matrice peuvent être utiles.

6° Méthode universelle ?

Quelle est *la* méthode qui peut permettre de

1. calculer l'inverse d'une matrice ;
2. trouver une base d'un sous-espace présenté par des équations ;
3. trouver des équations d'un sous-espace dont on connait une famille génératrice ;
4. calculer le rang d'une matrice ou d'une famille de vecteurs ?

7° Méthodes

Donner au moins deux méthodes pour

1. trouver la dimension d'un espace vectoriel ;
2. trouver l'inverse d'une matrice ;
3. montrer qu'une partie d'un ev est un sev ;
4. montrer que deux sous-espaces sont égaux ;
5. montrer qu'une famille est libre/génératrice/une base ;
6. montrer qu'une application linéaire est injective/surjective/bijective ;
7. montrer qu'une matrice est inversible.