MAT 253: Introduction à l'algèbre linéaire (II)

Chapitre V: Applications linéaires d'espaces vectoriels

On a étudié des propriétés élémentaires d'espaces vectoriels. Dans la plupart de cas, il est utile d'étudier la relation entre deux espaces vectoriels sur le même corps. Ceci sera donnée par les applications linéaires d'un espace vectoriel dans un autre espace vectoriel qui conservent la structure d'espace vectoriel.

Partout dans ce chapitre, on se fixe K un corps et E et F des espaces vectoriels sur K.

5.1. Applications linéaires

Tout d'abord, on rappelle la notion d'application d'ensembles. Soient X et Y deux ensembles. Une *application* T de X dans Y est une règle qui associe à chaque $x \in X$ un élément unique $y \in Y$, où y s'appelle l'*image* de x par T noté $y = T(x)$; et x, un *antécédent* de y. Une telle application est notée $T : X \to Y : x \mapsto T(x)$.

Soit $T : X \to Y$ une application. Alors on appelle l'*image* de X par T l'ensemble $\mathrm{Im}(T) = \{y \in Y \mid y = T(x)$ pour un certain $x \in X\}$. On dit que T est *injective* si tout $y \in Y$ admet au plus un antécédent; *surjective* si $\mathrm{Im}(T) = Y$; et *bijective* si T est à la fois injective et surjective. Par exemple, l'*application identité* $\mathbf{1}_X : X \to X : x \mapsto x$ est bijective. Si T est bijective, alors tout $y \in Y$ a exactement un antécédent noté $T^{-1}(y)$. Dans ce cas,

$$T^{-1} : Y \to X : y \mapsto T^{-1}(y)$$

est une application bijective appelée l'*inverse* de T. On dit que deux applications T_1 et T_2 de X dans Y sont *égales* si $T_1(x) = T_2(x)$ pour tout $x \in X$. Si $S : Y \to Z$ est une autre application, alors $ST : X \to Z : x \mapsto S(T(x))$ est une application de X dans Z, appelée le *composé* de T et S. La composition d'applications est associative. On voit aisément que $\mathbf{1}_Y T = T\mathbf{1}_X = T$ et $T^{-1}T = \mathbf{1}_X$ et $TT^{-1} = \mathbf{1}_Y$ lorsque $T : X \to Y$ est bijective.

Dès maintenant, on étudiera les applications entre des espaces vectoriels qui sont compatibles avec la structure d'espace vectoriel.

5.1.1. Définition. Une application $T : E \to F$ est dite *linéaire* si pour tous $u, v \in E$, $T(\alpha u) = \alpha T(u)$ et $T(u + v) = T(u) + T(v)$.

Remarque. Si $T : E \to F$ est linéaire, alors $T(0_E) = 0_F$.

Exemple. (1) L'application nulle $\mathbf{0} : E \to F : u \mapsto 0_F$ et l'application identitée $\mathbf{1}_E : E \to E : u \mapsto u$ sont linéaires.

(2) Soit $D^\infty[a, b]$ l'espace vectoriel réel des fonctions infiniment différentiables définies sur l'intervalle $[a, b]$. L'opérateur différentiel suivant est linéaire:

$$\partial : D^\infty[a, b] \to D^\infty[a, b] : f(t) \mapsto \frac{df(t)}{dt}.$$

(3) L'application suivante est linéaire:

$$T : M_{m\times n}(K) \to M_{n\times m}(K) : A \mapsto A^T.$$

(4) L'application

$$\sin x : \ \mathbb{R} \to \mathbb{R} : t \mapsto \sin t$$

n'est pas linéaire.

(5) Considérons le $\mathbb{R}$-espace $E = \mathbb{R}^2$ et le $\mathbb{C}$-espace $F = \mathbb{C}^2$. L'application

$$T : \mathbb{R}^2 \to \mathbb{C}^2 : (a, b) \mapsto (a, b)$$

n'est pas linéaire parce que E et F ne sont pas d'espaces vectoriels sur le même corps.

5.1.2. Proposition. Une application $T : E \to F$ est linéaire si, seulement si, pour tous $u, v \in E$, et $\alpha, \beta \in K$, on a $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$.

Démonstration. Supposons que T est linéaire. Alors pour tous $u, v \in E, \alpha, \beta \in K$, $T(\alpha u + \beta v) = T(\alpha u) + T(\beta v) = \alpha T(u) + \beta T(v)$. Réciproquement supposons que T satisfait à la condition. Alors $T(u+v) = T(1 \cdot u) + T(1 \cdot v) = 1 \cdot T(u) + 1 \cdot T(v)$ et

$$T(\alpha u) = T(\alpha u + 0 \cdot 0_E) = \alpha T(u) + 0 \cdot T(0_E) = \alpha T(u).$$

Ainsi T est linéaire. Ce qui achève la démonstration.

Exemple. Pour toute $A \in M_{m\times n}(K)$, les applications suivantes sont linéaires:

$$T_A : K^{(n)} \to K^{(m)} : v \mapsto Av \quad \text{et} \quad S_A : K^m \to K^n : u \mapsto uA.$$

En effet, pour tous $u, v \in K^m$ et $\alpha, \beta \in K$, on a $(\alpha u + \beta v)A = \alpha(uA) + \beta(vA)$.

On déduit par récurrence le résultat suivant de la proposition 5.1.2.

5.1.3. Corollaire. Si $T : E \to F$ est linéaire, alors

$$T(\alpha_1 u_1 + \cdots + \alpha_n u_n) = \alpha_1 T(u_1) + \cdots + \alpha_n T(u_n)$$

pour tous $u_i \in E$ et $\alpha_i \in K, \ i = 1, \ldots, n$.

5.1.4. Proposition. Soit E de dimension finie ayant pour base $\{u_1, \cdots, u_n\}$. Alors pour tous $v_1, \ldots, v_n \in F$, il existe une application linéaire unique $T : E \to F$ telle que $T(u_i) = v_i, i = 1, \ldots, n$.

Démonstration. Tout $u \in E$ s'écrit uniquement $u = \alpha_1 u_1 + \cdots + \alpha_n u_n, \ \alpha_i \in K$. Définissons une application $T : E \to F$ par $T(\alpha_1 u_1 + \cdots + \alpha_n u_n) = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Alors T est linéaire telle que $T(u_i) = T(1 \cdot u_i) = 1 \cdot v_i = v_i, i = 1, \ldots, n$. En outre si $S : E \to F$ est linéaire telle que $S(u_i) = v_i$, pour tout $1 \le i \le n$. Alors pour tout $u = \sum_{i=1}^n \alpha_i u_i \in E$, $S(u) = \sum_{i=1}^n \alpha_i S(u_i) = \sum_{i=1}^n \alpha_i T(u_i) = T(u)$. Donc $S = T$.

Remarque. Soit $\{u_1, \ldots, u_n\}$ une base de E. Si $S, T : E \to F$ sont des applications linéaires, d'après la proposition 5.1.4, on voit que $S = T$ si et seulement si $S(u_i) = T(u_i)$, $i = 1, \ldots, n$.

On désignera par $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F. On montrera que $\mathcal{L}(E, F)$ est un K-espace vectoriel.

5.1.5. Proposition. Soient $T, S \in \mathcal{L}(E, F)$ et $\alpha \in K$. Les applications suivantes sont linéaires:

$$\alpha T : E \to F : u \mapsto \alpha T(u) \quad \text{et} \quad T + S : E \to F : u \mapsto T(u) + S(u).$$

Démonstration. Pour tous $u, v \in E, \beta, \gamma \in K$,

$$\begin{array}{rcl} (\alpha T)(\beta u + \gamma v) & = & \alpha T(\beta u + \gamma v) = \alpha[\beta T(u) + \gamma T(v)] \\ & = & \beta \alpha T(u) + \gamma \alpha T(v) = \beta(\alpha T)(u) + \gamma(\alpha T)(u) \\ (T + S)(\beta u + \gamma v) & = & T(\beta u + \gamma v) + S(\beta u + \gamma v) \\ & = & \beta T(u) + \gamma T(v) + \beta S(u) + \gamma S(v) \\ & = & \beta[T(u) + S(u)] + \gamma[T(v) + S(v)] \\ & = & \beta(T + S)(u) + \gamma(T + S)(v). \end{array}$$

Donc αT et $T + S$ sont linéaires. Ceci achève la démonstration.

5.1.6. Lemme. Soient $T, S, T' \in \mathcal{L}(E, F)$ et $\alpha, \beta \in K$. Alors

(1)	$T + S = S + T$;	(2)	$(T + S) + T' = T + (S + T')$;
(3)	$1_K T = T$;	(4)	$\alpha(\beta T) = (\alpha\beta)T$;
(5)	$\alpha(T + S) = \alpha T + \alpha S$;	(6)	$(\alpha + \beta)T = \alpha T + \beta T$.

Démonstration. La démonstration est une vérification de routine. À titre d'exemple, on montrera (5). Pour tout $u \in E$,

$$\begin{array}{rcl} (\alpha(T + S))(u) & = & \alpha((T + S)(u)) \\ & = & \alpha(T(u) + S(u)) = \alpha T(u) + \alpha S(u) \\ & = & (\alpha T)(u) + (\alpha S)(u) \\ & = & (\alpha T + \alpha S)(u). \end{array}$$

Ainsi $\alpha(T + S) = \alpha T + \alpha S$. Ceci achève la démonstration.

5.1.7. Théorème. Muni des opérations définies dans la Proposition 5.1.5, $\mathcal{L}(E, F)$ est un K-espace vectoriel.

Démonstration. D'abord $0 : E \to F : u \mapsto 0_F$ est linéaire. Ainsi $\mathcal{L}(E, F)$ est non vide. Pour tout $T \in \mathcal{L}(E, F)$, l'application $-T : E \to F : u \mapsto -T(u)$ est linéaire telle que $T + (-T) = 0$. Il suit du lemme 5.1.6 que les autres axiomes de K-espace vectoriel sont aussi satisfaits. Ceci achève la démonstration.

5.1.8. Proposition. Soient $T : E \to F$ et $S : F \to G$ des applications linéaires de K-espaces vectoriels. Alors ST est linéaire. En outre,

(1) Pour tout $T' \in \mathcal{L}(E, F)$, $S(T + T') = ST + ST'$.

(2) Pour tout $S' \in \mathcal{L}(F, G)$, $(S + S')T = ST + S'T$.

Démonstration. Pour tous $u, v \in E, \alpha, \beta \in K$,

$$\begin{array}{rcl} (ST)(\alpha u + \beta v) & = & S[T(\alpha u + \beta v)] = S[\alpha T(u) + \beta T(v)] \\ & = & \alpha S(T(u)) + \beta S(T(v)) = \alpha(ST)(u) + \beta(ST)(v). \end{array}$$

Donc ST est linéaire. En outre pour tout $u \in E$,

$$\begin{array}{rcl} (S(T + T'))(u) & = & S((T + T')(u)) \\ & = & S(T(u) + T'(u)) \\ & = & S(T(u)) + S(T'(u)) \\ & = & (ST)(u) + (ST')(u). \end{array}$$

Ainsi $S(T + T') = ST + ST'$. De même, $(S + S')T = ST + S'T$. Ceci achève la dḿonstration.

5.2. Isomorphismes

On étudiera premièrement plus de propriétés d'une application linéaire d'espaces vectoriels et ensuite on étudiera quand deux espaces vectoriels sont éssentiellement identiques. On commence par deux sous-espaces associés à une application linéaire $T : E \to F$. Si $\mathcal{U}$ est une famille de vecteurs de E, on note $T(\mathcal{U}) = \{T(u) \mid u \in \mathcal{U}\}$, une famille de vecteurs de F.

5.2.1. Proposition. Soit $T \in (E, F)$. Alors
(1) $\text{Im}(T) = \{v \in F \,|\, v = F(u)$ pour un certain $u \in E\}$ est un sous-espace de F.
(2) $\text{Ker}(T) = \{u \in E \,|\, T(u) = 0_F\}$ est un sous-espace de E, appelé *noyau* de T.

Démonstration. (1) Soient $v_1, v_2 \in \text{Im}\,(T)$ et $\alpha_1, \alpha_2 \in K$. Alors $v_i = T(u_i)$ avec $u_i \in E,\ i = 1, 2$. Ainsi $\alpha_1 v_1 + \alpha_2 v_2 = T(\alpha_1 u_1 + \alpha_2 u_2) \in \text{Im}\,(T)$. Donc $\text{Im}(T)$ est un sous-espace de F.

(2) Soient $u, v \in \text{Ker}\,(T)$ et $\alpha, \beta \in K$. Alors $T(u) = T(v) = 0_F$. Or $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v) = 0_F$. Ainsi $\alpha u + \beta v \in \text{Ker}\,(T)$. Ce qui implique que $\text{Ker}(T)$ est un sous-espace de E.

On étudiera comment déterminer l'image d'une application linéaire.

5.2.2. Proposition. Soit $E =< \mathcal{U} >$ avec $\mathcal{U}$ une famille de vecteurs de E. Si $T \in \mathcal{L}(E, F)$, alors $\text{Im}(T) =< T(\mathcal{U}) >$. Par conséquent, T est surjective si et seulement si $F =< T(\mathcal{U}) >$.

Démonstration. Comme $T(\mathcal{U}) \subseteq \text{Im}(T)$, on a $< T(\mathcal{U}) >\subseteq \text{Im}(T)$. D'autre part, si $v \in \text{Im}(T)$, alors $v = T(w)$ avec $w \in E$. Or w s'écrit $w = \alpha_1 u_1 + \cdots + \alpha_n u_n,\ u_i \in \mathcal{U},\ \alpha_i \in K$. Donc $v = T(w) = \alpha_1 T(u_1) + \cdots + \alpha_n T(u_n) \in< T(\mathcal{U}) >$. Ce qui montre $\text{Im}(T) \subseteq< T(\mathcal{U}) >$. Ainsi $\text{Im}(T) =< T(\mathcal{U}) >$. Ceci achève la démonstration.

Le résultat suivant dit que l'image d'une famille liée par une application linéaire est toujours liée.

5.2.3. Lemme. Soient $T \in \mathcal{L}(E, F)$ et $\mathcal{U}$ une famille de vecteurs de E.
(1) Si $T(\mathcal{U})$ est libre, alors $\mathcal{U}$ est libre.
(2) Si T est injective, alors $\mathcal{U}$ est libre si et seulement si $T(\mathcal{U})$ est libre.

Démonstration. (1) Supposons que $T(\mathcal{U})$ est libre. Si $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0_E,\ u_i \in \mathcal{U}, \alpha_i \in K$, alors $\alpha_1 T(u_1) + \cdots + \alpha_n T(u_n) = 0_F$. Comme $\{T(u_1), \ldots, T(u_n)\}$ est libre, $\alpha_1 = \cdots = \alpha_n = 0$. Donc $\mathcal{U}$ est libre.

(2) Supposons que T est injective et $\mathcal{U}$ est libre. Si $\alpha_1 T(u_1) + \cdots + \alpha_n T(u_n) = 0_F, \alpha_i \in K, u_i \in \mathcal{U}$, alors $T(\alpha_1 u_1 + \cdots + \alpha_n u_n) = T(0_E)$. Il suit de l'injectivité de T que $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0_E$. Comme $\{u_1, \ldots, u_n\}$ est libre, $\alpha_1 = \cdots = \alpha_n = 0$. Ceci montre que $T(\mathcal{U})$ est libre. Le preuve s'achève.

On donne maintenant un critère pour qu'une application linéaire soit injective.

5.2.4. Proposition. Soient $T \in \mathcal{L}(E, F)$ et $\mathcal{B}$ une base de E. Alors les conditions suivantes sont équivalentes:
(1) T est injective.
(2) $\text{Ker}(T) = 0$.
(3) La famille $T(\mathcal{B})$ est libre.

Démonstration. Supposons que T est injective. Si $u \in\text{Ker}(T)$, alors $T(u) = 0_F = T(0_E)$. Ainsi $u = 0_E$ d'après l'injectivité de T. Par conséquent, $\text{Ker}(T) = 0$.

Supposons que $\text{Ker}(T) = 0$. Soient $v_1, \ldots, v_r \in T(\mathcal{B})$, c'est-à-dire, $v_i = T(u_i)$ avec $u_i \in \mathcal{B},\ i = 1, \ldots, r$. Si $\alpha_1 v_1 + \cdots + \alpha_r v_r = 0_F$, $\alpha_i \in K$, alors $T(\alpha_1 u_1 + \cdots + \alpha_r u_r) = 0_F$, c'est-à-dire, $\alpha_1 u_1 + \cdots + \alpha_r u_r \in \text{Ker}(T)$.

Comme $\mathrm{Ker}(T) = \{0_E\}$ par l'hypothèse, $\alpha_1 u_1 + \cdots + \alpha_r u_r = 0_E$. Comme $\mathcal{B}$ est libre, $\alpha_i = 0$, $i = 1, \ldots, r$. Ceci montre que $T(\mathcal{B})$ est libre.

Supposons enfin que $T(\mathcal{B})$ est libre. Soient $u, v \in E$ tels que $T(u) = T(v)$. Écrivons $u = \sum_{i=1}^n \alpha_i u_i, v = \sum_{i=1}^n \beta_i u_i,\ \alpha_i, \beta_i \in K, u_i \in \mathcal{B}$. Alors

$$\sum_{i=1}^n (\alpha_i - \beta_i) T(u_i) = \sum_{i=1}^n \alpha_i T(u_i) - \sum_{i=1}^n \beta_i T(u_i) = T(u) - T(v) = 0_F.$$

Comme $\{T(u_1), \ldots, T(u_n)\}$ est libre, on a $\alpha_i - \beta_i = 0$, $i = 1, \ldots, n$. Ainsi $u = v$. Ceci montre que T est injective. Ceci achève la démonstration.

Remarque. Il suit de la proposition 5.2.4 que $\dim(E) \le \dim(F)$ lorsqu'il existe une application linéaire injective $T : E \to F$.

5.2.5. Définition. On dit que $T \in \mathcal{L}(E, F)$ est un *isomorphisme* si T est bijective.

Exemple. (1) L'application identitée $\mathbb{1} : E \to E : u \mapsto u$ est un isomorphisme.

(2) L'application nulle $\mathbf{0} : E \to F$ est un isomorphisme si et seulement si E et F sont tous nuls.

(3) L'application $T : M_{m\times n}(K) \to M_{n\times m}(K) : A \mapsto A^T$ est un isomorphisme. En particulier, l'application linéaire suivante est un isomorphisme:

$$S : K^n \to K^{(n)} : (a_1, \ldots, a_n) \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

5.2.6. Proposition. Soient E, F et G des K-espace vectoriels.

(1) Si $T : E \to F$ est un isomorphisme, alors $T^{-1} : F \to G$ est un isomorphisme.

(2) Si $T : E \to F$ et $S : F \to G$ sont des isomorphismes, alors $ST : E \to G$ est un isomorphism et $(ST)^{-1} = T^{-1}S^{-1}$.

Démonstration. (1) Supposons que $T : E \to F$ est un isomorphisme. Par définition, T est bijective. Ainsi $T^{-1} : F \to E$ est une bijection. On veut montrer que T^{-1} est linéaire. Soient $v_1, v_2 \in F$ et $\alpha_1, \alpha_2 \in K$. Posons $u_i = T^{-1}(v_i) \in E,\ i = 1, 2$. D'après la définition, $T(u_i) = v_i, i = 1, 2$. Come $T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T(u_1) + \alpha_2 T(u_2) = \alpha_1 v_1 + \alpha_2 v_2$, on a

$$T^{-1}(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 u_1 + \alpha_2 u_2 = \alpha_1 T^{-1}(v_1) + \alpha_2 T^{-1}(v_2).$$

D'où, T^{-1} est linéaire.

(2) Soient $T : E \to F$ et $S : F \to G$ des isomorphismes. Alors $ST : E \to G$ est une bijection qui est linéaire d'après la proposition 5.1.8. C'est-à-dire, ST est un isomorphisme. Enfin, on vérifie aisément que $(ST)(T^{-1}S^{-1}) = \mathbf{1}_G$ et $(T^{-1}S^{-1})(ST) = \mathbf{1}_E$. D'où, $(ST)^{-1} = T^{-1}S^{-1}$. Ceci achève la démonstration.

5.2.7. Corollaire. Une application linéaire $T : E \to F$ est un isomorphisme si et seulement s'il existe une application linéaire $S : F \to E$ telle que $ST = \mathbf{1}_E$ et $TS = \mathbf{1}_F$.

Démonstration. Supposons que T est un isomorphisme. D'après la proposition 5.2.6, $T^{-1} : F \to E$ est une application linéaire telle que $T^{-1}T = \mathbf{1}_E$ et $TT^{-1} = \mathbf{1}_F$.

Supposons réciproquement qu'il existe une application linéaire $S : F \to E$ telle que $ST = \mathbf{1}_E$ et $TS = \mathbf{1}_F$. Pour tout $v \in F$, $S(v) \in E$ est tel que $T(S(v)) = (TS)(v) = \mathbf{1}_F(v) = v$. Ce qui montre que T est surjective.

Si $T(u_1) = T(u_2)$ avec $u_i \in E,\ i = 1, 2$, alors $S(T(u_1)) = S(T(u_2))$, c'est-à-dire, $(ST)(u_1) = (ST)(u_2)$. Donc $u_1 = u_2$ comme $ST = \mathbf{1}_E$. Ceci implique que T est injective. Par conséquent, T est un isomorphisme. Ceci achève la démonstration.

5.2.8. Proposition. Si $\{u_1, \ldots, u_n\}$ est une base de E, alors l'application

$$T : K^{(n)} \to E : \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \mapsto (u_1, \ldots, u_n) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

est un isomorphisme.

Démonstration. Il est évident que T est linéaire. D'après le lemme 4.3.2(3), l'application

$$S : E \to K^{(n)} : u \mapsto P^{\{u\}}_{\{u_1, \cdots, u\}}.$$

est linéaire. On vérifie aisément que $ST = \mathbf{1}_{K^{(n)}}$ et $TS = \mathbf{1}_E$. D'après le corollaire 5.2.7, T est un isomorphisme. Ceci achève la démonstration.

5.2.9. Proposition. Soit $\mathcal{B}$ une base de E. Si $T \in \mathcal{L}(E, F)$, alors T est un isomorphisme $\Leftrightarrow T(\mathcal{B})$ est une base de F.

Démonstration. T est un isomorphisme si et seulement si T est injective et surjective si et seulement si $T(\mathcal{B})$ est libre et $F =< T(\mathcal{B}) >$ si et seulement si $T(\mathcal{B})$ est une base F. Ceci achève la démonstration.

5.2.10. Corollaire. Soit $T \in \mathcal{L}(E, F)$ un isomorphisme. Si $\mathcal{U}$ est une famille de vecteurs de E, alors $\mathcal{B}$ est une base de E si et seulement si $T(\mathcal{B})$ est une base de F.

Démonstration. Si $\mathcal{B}$ est une base de E, alors $T(\mathcal{B})$ est une base de F d'après la proposition 5.2.9.

Réciproquement supposons que $T(\mathcal{B})$ est une base de F. Alors $\mathcal{B}$ est libre car $T(\mathcal{B})$ est libre. Ensuite pour tout $u \in E$, on a $T(u) = \alpha_1 T(u_1) + \cdots \alpha_n T(u_n)$ avec $u_i \in \mathcal{B},\ \alpha_i \in K$ car $F =< T(\mathcal{B}) >$. Ainsi $T(u) = T(\alpha_1 u_1 + \cdots + \alpha_n u_n)$. Ce qui donne $u = \alpha_1 u_1 + \cdots + \alpha_n u_n$ car T est injective. Par conséquent, $\mathcal{B}$ est une base de E. Ceci achève la démonstration.

On dit que E et F sont *isomorphes*, noté $E \cong F$, s'il existe un isomorphisme $T : E \to F$. Dans ce cas, on peut identifier E avec F en identifiant $u \in E$ avec $T(u) \in F$.

Exemple. (1) On a vu que $K^{(n)} \cong K^n$, pour tout $n \geq 1$.

(2) On sait que $\{0_K\}$ est un K-espace vectoriel nul. Si E est un K-espace vectoriel nul, alors $E \cong \{0_K\}$.

5.2.11. Proposition. La relation d'isomorphisme de K-espaces vectoriels est une relation d'équivalence.

Démonstration. D'abord, comme $\mathbf{1}_E : E \to E$ est un isomorphisme, on a $E \cong E$. Ensuite, si $E \cong F$, alors il existe un isomorphisme $T : E \to F$. Comme $T^{-1} : F \to E$ est également un isomorphisme, on a $F \cong E$. Enfin, si $E \cong F$ et $F \cong G$, alors il existent des isomorphismes $T : E \to F$ et $S : F \to G$. D'après le corollaire 5.2.7, $ST : E \to G$ est un isomoprhisme, et donc $E \cong G$. Ceci achève la démonstration.

Il suit du corollaire 5.2.9 que deux K-espaces vectoriels isomorphes ont la même dimension. La réciproque est vraie si l'un de ces deux espaces est de dimension finie.

5.2.12. Théorème. Soit E, F des K-espaces vectoriels. Si E est de dimension finie, alors $E \cong F$ si et seulement si $\dim(E) = \dim(F)$.

Démonstration. La necessité suit immédiatement du corollaire 5.2.9. Supposons que $\dim(E) = \dim(F) = n$. Alors $E \cong K^{(n)}$ et $F \cong K^{(n)}$. Ainsi $E \cong F$. Ceci achève la démonstration.

Remarque. Le résultat ci-dessus nous dit qu'un espace vectoriel de dimension finie est unique déterminé (à isomorphisme près) par sa dimension. Par conséquent, pour tout entier $n \geq 1$, le K-espace vectoriel K^n est essentiellement le seul de dimension n.

Exemple. ${}_{\mathbb{R}}\mathbb{C} \cong \mathbb{R}^2$ et $K_n[x] \cong K^{(n)}$, pour tout $n \geq 1$.

5.3. Représentation matricielle d'applications linéaires

Cette section a pour but de représenter les applications linéaires d'espaces vectoriels de dimension finie par les matrices. Ce qui nous permet d'étudier les applications linéaires en appliquant les propriétés de matrices.

Partout dans cette section, on se fixe E et F deux K-espace vectoriels de dimension finie.

5.3.1. Définition. Soient $\{u_1, \ldots, u_n\}$ une base de E, et $\{v_1, \ldots, v_m\}$ une base de F. Si $T \in \mathcal{L}(E, F)$, alors $\{T(u_1), \ldots, T(u_n)\}$ est une famille de vecteurs de F. On appelle $P_{\{v_1,\ldots,v_m\}}^{\{T(u_1),\ldots,T(u_n)\}}$ la *matrice* de T dans les bases $\{u_1, \ldots, u_n\}$ et $\{v_1, \ldots, v_m\}$, notée $[T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}$. Remarquons que

$$(T(u_1), \ldots, T(u_n)) = (v_1, \ldots, v_m)[T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}.$$

Exemple. (1) $[\mathbf{0}]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}} = 0_{m\times n}$ et $[\mathbf{1}_E]_{\{u_1,\ldots,u_n\}}^{\{u_1,\ldots,u_n\}} = I_n$.

(2) Une matrice $A \in M_{m\times n}(K)$ définit deux applications linéaires

$$L_A : K^{(n)} \to K^{(m)} : v \mapsto Av; \quad R_A : K^m \to K^n : u \mapsto uA.$$

On voit que la matrice de L_A dans les bases canoniques est A, mais celle-ci de R_A est A^T.

Le résultat suivant dit qu'une application linéaire est uniquement déterminée par sa matrice.

5.3.2. Proposition. Soient $\{u_1, \ldots, u_n\}$ une base de E et $\{v_1, \ldots, v_m\}$ une base de F. Si $T, S \in \mathcal{L}(E, F)$, alors $T = S$ si et seulement si

$$[T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}} = [S]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}.$$

Démonstration. Il suffit de montrer la suffisance. Si $[T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}} = [S]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}$, alors

$$(T(u_1), \ldots, T(u_n)) = (S(u_1), \ldots, S(u_n)),$$

c'est-à-dire, $T(u_i) = S(u_i)$ $i = 1, \ldots, n$. D'après la proposition 5.1.4, $T = S$. Ceci achève la démonstration.

Le résultat suivant nous permet de déterminer si une application d'espaces vectoriels de dimension finie est linéaire ou non.

5.3.3. Proposition. Soient $\{u_1, \ldots, u_n\}$ une base de E et $\{v_1, \ldots, v_m\}$ une base de F. Une application $T : E \to F$ est linéaire si et seulement s'il existe une matrice $A = (a_{ij})_{m\times n}$ sur K telle que, pour tous $x_1, \ldots, x_n \in K$, on a

$$T(x_1u_1 + \cdots + x_nu_n) = (a_{11}x_1 + \cdots + a_{1n}x_n)v_1 + \cdots + (a_{m1}x_1 + \cdots + a_{mn}x_n)v_m,$$

c'est-à-dire, $P^{\{T(u)\}}_{\{v_1,\ldots,v_m\}} = AP^{\{u\}}_{\{u_1,\ldots,u_n\}}$, pour tout $u \in E$. Dans ce cas, $A = [T]^{\{u_1,\ldots,u_n\}}_{\{v_1,\ldots,v_m\}}$.

Démonstration. On commence par la necessité. Supposons que T est linéaire et posons $A = (a_{ij})_{m\times n} = [T]^{\{u_1,\ldots,u_n\}}_{\{v_1,\ldots,v_m\}}$. Par définition, $(T(u_1), \ldots, T(u_n)) = (v_1, \ldots, v_m)A$, c'est-à-dire, $T(u_j) = \sum_{i=1}^m a_{ij}v_i, \ j = 1, \ldots, n$. Si $u = \sum_{j=1}^n x_ju_j \in E$, alors

$$\begin{array}{rcl} T(u) & = & \sum_{j=1}^n x_j\, T(u_j) = \sum_{j=1}^n x_j(\sum_{i=1}^m a_{ij}v_i) \\ & = & \sum_{j=1}^n \sum_{i=1}^m a_{ij}x_jv_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}x_j\right) v_i\,. \end{array}$$

Supposons réciproquement que T satisfait à la condition énoncée dans la proposition. Si $u = \sum_{j=1}^n x_ju_j, \ v = \sum_{i=1}^n y_ju_j \in E$ et $\alpha, \beta \in K$, alors $\alpha u + \beta v = \sum_{j=1}^n(\alpha x_j + \beta y_j)u_j$ et donc

$$\begin{array}{rcl} T(\alpha u + \beta v) & = & \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}(\alpha x_j + \beta y_j)\right) v_i \\ & = & \alpha \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}x_j\right) v_i + \beta \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}y_j\right) v_i \\ & = & \alpha\, T(u) + \beta\, T(v)\,. \end{array}$$

Ainsi T est linéaire. De plus, la colonne de u_j dans $\{u_1, \ldots, u_n\}$ est $e_j \in K^{(n)}$. Ainsi la colonne de $T(u_j)$ dans $\{v_1, \ldots, v_m\}$ est Ae_j, la j-ième colonne de A. Ceci implique que A est la matrice de $\{T(u_1), \ldots, T(u_n)\}$ dans $\{v_1, \ldots, v_m\}$, c'est-à-dire, la matrice de T dans les bases $\{u_1, \ldots, u_n\}$ et $\{v_1, \ldots, v_m\}$. Ceci achève la démonstration.

Remarque. Si $T : E \to F$ est linéaire, alors T est donnée par

$$T : E \to F : (u_1, \ldots, u_n)B \mapsto (v_1, \ldots, v_m)AB.$$

Exemple. Considérons le plan $\mathbb{R}^2$. On se fixe un angle θ. La rotation $\rho_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ d'angle θ est linéaire. En effet, soit $\{e_1, e_2\}$ la base canonique. Pour tout $u = (x, y) \in \mathbb{R}^2$, on a $u = xe_1 + ye_2 = (r\cos\phi, r\sin\phi)$, où $r \geq 0$ et ϕ est un angle. D'après la définition,

$$\begin{array}{rcl} \rho_\theta(u) & = & \rho_\theta(r\cos\phi, r\sin\phi) \\ & = & (r\cos(\theta+\phi),\ r\sin(\theta+\phi)) \\ & = & (r\cos\theta\cos\phi - r\sin\theta\sin\phi,\ r\sin\theta\cos\phi - r\cos\theta\sin\phi) \\ & = & [(\cos\theta)x - (\sin\theta)y]e_1 + [(\sin\theta)x - (\cos\theta)y]e_2. \end{array}$$

D'où ρ_θ est linéaire, et

$$[\rho_\theta]^{\{e_1,e_2\}}_{\{e_1,e_2\}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Le résultat suivant nous dit comment trouver l'image d'une application linéaire à l'aide de sa matrice.

5.3.4. Théorème. Soient $\{u_1, \ldots, u_n\}$ une base de E et $\{v_1, \ldots, v_m\}$ une base de F. Soient $T \in \mathcal{L}(E, F)$ et $A = [T]^{\{u_1,\ldots,u_n\}}_{\{v_1,\ldots,v_m\}}$.

(1) Un vecteur $v \in F$ appartient à $\mathrm{Im}(T)$ si et seulement si $P_{\{v_1,\ldots,v_m\}}^{\{v\}}$ appartient à l'espace-colonne $\mathcal{C}(A)$ de A.

(2) Si $\{B_1, \ldots, B_r\}$ est une base de $\mathcal{C}(A)$ et $w_i = (u_1, \ldots, u_n)B_i$, $i = 1, \cdots, r$, alors $\{w_1, \ldots, w_r\}$ est une base de $\mathrm{Im}(T)$. En particulier, $\dim(\mathrm{Im}(T)) = \mathrm{rg}(A)$.

Démonstration. Soit $A = (A_1 \cdots A_n)$, où $A_j = P_{\{v_1,\ldots,v_m\}}^{\{T(u_j)\}}$, $j = 1, \ldots, n$. Comme $E =< u_1, \ldots, u_n >$, d'après la proposition 5.2.2, $\mathrm{Im}(T) =< T(u_1), \ldots, T(u_n) >$. Pour tout $v \in F$, on voit que $v \in \mathrm{Im}(T)$ si et seulement si v est combinaison linéaire de $T(u_1), \ldots, T(u_n)$ si et seulement si, d'après le lemme 4.3.3(2), $P_{\{v_1,\ldots,v_m\}}^{\{v\}}$ est combinaison linéaire de $A_1, \ldots, A_n$, c'est-à-dire, $P_{\{v_1,\ldots,v_m\}}^{\{v\}} \in \mathcal{C}(A)$. Ceci montre (1).

Comme $\{v_1, \cdots, v_m\}$ est une base de F, on a un isomorphism

$$\Phi : K^{(m)} \to F : B \mapsto (v_1, \cdots, v_m)B.$$

D'après la partie (1), restreindre Φ à $\mathcal{C}(A)$ nous donne un isomorphisme

$$\Psi : \mathcal{C}(A) \to \mathrm{Im}(T) : B \mapsto (v_1, \cdots, v_m)B$$

tel que $\Psi(B_j) = w_j$, $j = 1, \cdots, r$. Comme $\{B_1, \ldots, B_r\}$ est une base de $\mathrm{C}(A)$, d'après le corollaire 5.2.10, $\{w_1, \cdots, w_r\}$ est une base de $\mathrm{Im}(T)$. Ceci achève la démonstration.

Remarque. On appelle *rang* de T la dimension de $\mathrm{Im}(T)$, et noté $\mathrm{rg}(T)$.

Le résultat suivant nous dit comment trouver le noyau d'une application linéaire à l'aide de sa matrice.

5.3.5. Théorème. Soient $\{u_1, \ldots, u_n\}$ une base de E et $\{v_1, \ldots, v_m\}$ une base de F. Soient $T \in \mathcal{L}(E, F)$ et $A = [T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}$.

(1) Un vecteur $u \in E$ appartient à $\mathrm{Ker}(T)$ si et seulement si $P_{\{u_1,\ldots,u_n\}}^{\{u\}}$ appartient au noyau $\mathcal{N}(A)$ de A.

(2) Si $\{B_1, \ldots, B_s\}$ est une base de $\mathcal{N}(A)$ et $w_j = (u_1, \ldots, u_n)B_j$, $j = 1, \ldots, s$, alors $\{w_1, \ldots, w_s\}$ est une base de $\mathrm{Ker}(T)$. En particulier, $\dim(\mathrm{Ker}(T)) = \dim\mathcal{N}(A)$.

Démonstration. (1) Par définition, $u \in \mathrm{Ker}(T)$ si et seulement si $T(u) = 0_F$ si et seulement si $P_{\{v_1,\ldots,v_m\}}^{\{T(u)\}} = 0_{m\times 1}$ si et seulement si, d'après la proposition 5.3.3 $AP_{\{u_1,\ldots,u_n\}}^{\{u\}} = 0_{m\times 0}$, c'est-à-dire, $P_{\{u_1,\ldots,u_n\}}^{\{u\}} \in \mathcal{N}(A)$.

(2) Comme $\{u_1, \cdots, u_n\}$ est une base de E, on a un isomorphisme

$$\Phi : K^{(n)} \to E : B \mapsto (u_1, \cdots, u_n)B.$$

D'après la partie (1), restreindre Φ à $\mathcal{N}(A)$ nous donne un isomorphisme

$$\Psi : \mathcal{N}(A) \to \mathrm{Ker}(T) : B \mapsto (v_1, \ldots, v_m)B$$

tel que $\Phi(B_j) = w_j$, $j = 1, \ldots, s$. Comme $\{B_1, \ldots, B_s\}$ est une base de $\mathcal{N}(A)$, d'après le corollaire 5.2.10, $\{w_1, \ldots, w_s\}$ est une base de $\mathrm{Ker}(T)$. Ceci achève la démonstration.

Le résultat suivant donne le lien enter les dimensions de l'image et du noyau d'une application linéaire.

5.3.6. Corollaire. Soit $T \in \mathcal{L}(E, F)$. Alors $\dim(E) = \dim(\mathrm{Im}(T)) + \dim(\mathrm{Ker}(T))$.

Démonstration. Soient $\{u_1, \ldots, u_n\}$ une base de E et $\{v_1, \ldots, v_m\}$ une base de F. Soit $A = [T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}$. D'après les théorèmes 5.3.4 et 5.3.5, $\dim(\mathrm{Im}(T)) = \mathrm{rg}(A)$ et $\dim(\mathrm{Ker}(T)) = \dim\mathcal{N}(A) = n - \mathrm{rg}(A)$, où la dernière égalité suit du théorème 4.8.1. Ainsi $\dim(\mathrm{Im}(T)) + \dim(\mathrm{Ker}(T)) = n = \dim(E)$. Ceci achève la démonstration.

Le résultat suivant nous donne une caractérisation de l'injectivité et de la surjectivité d'une application linéaire en terme du rang de sa matrice.

5.3.7. Proposition. Soient $\{u_1, \ldots, u_n\}$ une base de E et $\{v_1, \ldots, v_m\}$ une base de F. Soient $T \in \mathcal{L}(E, F)$ et $A = [T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}$.

(1) T est injective si et seulement si $\mathrm{rg}(A) = n$. Dans ce cas, $\dim(E) \leq \dim(F)$.

(2) T est surjective si et seulement si $\mathrm{rg}(A) = m$. Dans ce cas, $\dim(F) \leq \dim(E)$.

(3) T est un isomorphisme si et seulement si A est inversible.

Démonstration. (1) T est injective si et seulement si $\mathrm{Ker}(T) = 0$ si et seulement si $\dim(\mathrm{Ker}(T)) = 0$ si et seulement si $\dim(E) - \mathrm{rg}(A) = 0$, c'est-à-dire, $\mathrm{rg}(A) = n$.

(2) T est surjective si et seulement si $\mathrm{Im}(T) = F$ si et seulement si $\dim(\mathrm{Im}(T)) = \dim(F)$ si et seulement si $\mathrm{rg}(A) = m$.

(3) T est un isomorphisme si et seulement si T est injective et surjective si et seulement si $\mathrm{rg}(A) = n = m$ si et seulement si A est inversible. Ceci achève la démonstration.

5.3.8. Lemme. Soient $\{u_1, \ldots, u_n\}$ une base de E et $\{v_1, \ldots, v_m\}$ une base de F. Si $T, S \in \mathcal{L}(E, F)$ et $\alpha, \beta \in K$, alors

$$[\alpha T + \beta S]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}} = \alpha[T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}} + \beta[S]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}.$$

Démonstration. Posons $A = [T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}$ et $B = [S]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}$. Alors $(T(u_1), \ldots, T(u_n)) = (v_1, \ldots, v_m)A$ et $(S(u_1), \ldots, S(u_n)) = (v_1, \ldots, v_m)B$. Or

$$\begin{aligned}
& ((\alpha T + \beta S)(u_1), \ldots, (\alpha T + \beta S)(u_n)) \\
= \; & (\alpha T(u_1) + \beta S(u_1), \ldots, \alpha T(u_n) + \beta S(u_n)) \\
= \; & \alpha(T(u_1), \ldots, T(u_n)) + \beta(S(u_1), \ldots, S(u_n)) \\
= \; & \alpha(v_1, \ldots, v_m)A + \beta(v_1, \ldots, v_m)B \\
= \; & (v_1, \ldots, v_m)(\alpha A + \beta B).
\end{aligned}$$

Par conséquent, $[\alpha T + \beta S]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}} = \alpha A + \beta B$. Ceci achève la démonstration.

5.3.9. Lemme. Soient $\{u_1, \ldots, u_n\}$ une base de E et $\{v_1, \ldots, v_m\}$ une base de F. Pour toute $A \in M_{m\times n}(K)$, il existe une unique $T \in \mathcal{L}(E, F)$ telle que $[T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}} = A$.

Démonstration. Posons $(w_1, \ldots, w_n) = (v_1, \ldots, v_m)A$. D'après la proposition 5.1.4, il existe une unique $T \in \mathcal{L}(E, F)$ telle que $T(u_i) = w_i, i = 1, \ldots, n$. Or

$$(T(u_1), \ldots, T(u_n)) = (w_1, \ldots, w_n) = (v_1, \ldots, v_m)A.$$

Par conséquent, $[T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}} = A$. Ceci achève la démonstration.

5.3.10. Théorème. Soient $\dim E = n$ et $\dim F = m$. Alors $\mathcal{L}(E, F) \cong M_{m\times n}(K)$. En particulier, $\dim\mathcal{L}(E, F) = mn$.

Démonstration. On se fixe une base $\{u_1, \ldots, u_n\}$ de E et une base $\{v_1, \ldots, v_m\}$ de F. Considérons l'application suivante

$$\Phi : \mathcal{L}(E, F) \to M_{m\times n}(K) : T \mapsto [T]_{\{v_1,\ldots,v_m\}}^{\{u_1,\ldots,u_n\}}.$$

Alors Φ est linéaire d'après le lemme 5.3.8 et bijective d'après le lemme 5.3.9. Ainsi Φ est un isomorphisme. Ceci achève la démonstration.

5.3.11. Théorème. Soient E, F, G des K-espaces vectoriels ayant pour base $\{u_1, \ldots, u_n\}$, $\{v_1, \ldots, v_m\}$, $\{w_1, \ldots, w_p\}$, respectivement. Si $T \in \mathcal{L}(E, F)$ et $S \in \mathcal{L}(F, G)$, alors

$$[ST]^{\{u_1,\ldots,u_n\}}_{\{w_1,\ldots,w_p\}} = [S]^{\{v_1,\ldots,v_m\}}_{\{w_1,\ldots,w_p\}} \cdot [T]^{\{u_1,\ldots,u_n\}}_{\{v_1,\ldots,v_m\}}.$$

Démonstration. Posons $[S]^{\{v_1,\ldots,v_m\}}_{\{w_1,\ldots,w_p\}} = A$ et $[T]^{\{u_1,\ldots,u_n\}}_{\{v_1,\ldots,v_m\}} = B = (B_1 \cdots B_n)$, où $B_i = P^{\{T(u_i)\}}_{\{v_1,\ldots,v_m\}}$, $i = 1, \ldots, n$. Alors $P^{\{S(T(u_i))\}}_{\{w_1,\ldots,w_p\}} = AB_i$, $i = 1, \ldots, n$. Donc

$$[ST]^{\{u_1,\ldots,u_n\}}_{\{w_1,\ldots,w_p\}} = P^{\{S(T(u_1),\ldots,S(T(u_n))\}}_{\{w_1,\ldots,w_p\}} = (AB_1, \cdots, AB_n) = AB.$$

Ceci achève la démonstration.

5.3.12. Corollaire. Soit $T : E \to F$ un isomorphisme. Si $\{u_1, \ldots, u_n\}$ est une base de E et $\{v_1, \ldots, v_n\}$ est une base de F, alors

$$[T^{-1}]^{\{v_1,\ldots,v_n\}}_{\{u_1,\ldots,u_n\}} = \left([T]^{\{u_1,\ldots,u_n\}}_{\{v_1,\ldots,v_n\}}\right)^{-1}.$$

Démonstration. Posons $A = [T]^{\{u_1,\ldots,u_n\}}_{\{v_1,\ldots,v_n\}}$ et $B = [T^{-1}]^{\{v_1,\ldots,v_n\}}_{\{u_1,\ldots,u_n\}}$. Comme $T^{-1}T = \mathbf{1}_E$ et $[\mathbf{1}_E]^{\{u_1,\ldots,u_n\}}_{\{u_1,\ldots,u_n\}} = I_n$, on a $BA = I_n$ d'après le théorème 5.3.11. De même, $AB = I_n$. Ceci achève la démonstration.

5.4. Endomorphismes

Une application linéaire $T : E \to E$ s'appelle *endomorphisme* de E; et *automorphisme* de E si T est un isomorphisme. Par exemple, l'application nulle $\mathbf{0} : E \to E : u \mapsto 0_E$ est un endomorphisme de E, et l'application identité $\mathbf{1}_E$ de E est un automorphisme de E. Désignerons par $\mathrm{End}(E)$ l'ensemble des endomorphismes de E qui, d'après le théorème 5.1.7, est un espace vectoriel sur K. Le but de cette section est étudier, à l'aide de la théorie des matrices, les endomorphismes d'un espace vectoriel de dimension finie.

5.4.1. Théorème. Soit E de dimension finie. Alors les conditions suivantes sont équivalentes pour $T \in \mathrm{End}(E)$:

(1) T est un automorphisme.
(2) T est injective.
(3) T est surjective.

Démonstration. Comme E est de dimension finie, $\dim(E) = \dim(\mathrm{Ker}(T)) + \dim(\mathrm{Im}(T))$. Or T est injective si et seulement si $\dim(\mathrm{Ker}(T)) = 0$ si et seulement si $\dim(E) = \dim(\mathrm{Im}(T))$ si et seulement si $\mathrm{Im}(T) = E$ si et seulement si T est surjective. Ceci achève la démonstration.

5.4.2. Définition. Soit $\{u_1, \ldots, u_n\}$ une base de E. Pour $T \in \mathrm{End}\,(E)$, on définit la *matrice* de T dans $\{u_1, \ldots, u_n\}$, notée $[T]_{\{u_1,\ldots,u_n\}}$, comme étant la matrice de la famille $\{T(u_1), \ldots, T(u_n)\}$ dans cette base. Ainsi

$$(T(u_1), \ldots, T(u_n)) = (u_1, \ldots, u_n)[T]_{\{u_1,\ldots,u_n\}}.$$

Exemple. (1) $[\mathbf{0}_E]_{\{u_1,\ldots,u_n\}} = 0_{n\times n}$ et $[\mathbf{1}_E]_{\{u_1,\ldots,u_n\}} = I_n$.
(3) Soit $A \in M_{n\times n}(K)$. Alors la matrice de

$$L_A : K^{(n)} \to K^{(n)} : v \mapsto Av$$

dans la base canonique est A. Et la matrice de

$$R_A : K^n \to K^n : u \mapsto uA$$

dans la base canonique est A^T.

On voit que si $T \in \mathrm{End}(E)$, alors $T^2 = TT \in \mathrm{End}(E)$. En général on définit $T^0 = \mathbf{1}_E$ et $T^n = TT^{n-1}$ pour tout $n \geq 1$. Le résultat suivant suit de la proposition 5.3.3, le corollaire 5.3.7, la proposition 5.3.12 et le corollaire 5.3.13.

5.4.3. Théorèm. Soit $T \in \mathrm{End}(E)$. Si $\{u_1, \ldots, u_n\}$ est une base de E, alors
(1) $P^{\{T(u)\}}_{\{u_1,\ldots,u_n\}} = [T]_{\{u_1,\ldots,u_n\}} P^{\{u\}}_{\{u_1,\ldots,u_n\}}$, pour tout $u \in E$.
(2) T est un automorphisme si et seulement si $[T]_{\{u_1,\ldots,u_n\}}$ est inversible.
(3) $[T^r]_{\{u_1,\ldots,u_n\}} = \left([T]_{\{u_1,\ldots,u_n\}}\right)^r$, pour tout $r \geq 0$.
(4) $[T^r]_{\{u_1,\ldots,u_n\}} = \left([T]_{\{u_1,\ldots,u_n\}}\right)^r$, pour tout entier r lorsque T est un automorphisme.

Exemple. On voit aisément que la rotation ρ_θ du plan d'angle θ est un automorphisme et $\rho_\theta^n = \rho_{n\theta}$, pour tout entier n. Or

$$[\rho_\theta]_{\{e_1,e_2\}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

et

$$[\rho_{n\theta}]_{\{e_1,e_2\}} = \begin{pmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix}.$$

D'après le théorème 5.4.3(4), on a

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^n = \begin{pmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix}, \text{ pour tout } n \in \mathbb{Z}.$$

Le résultat suivant nous dit comment sont reliées les matrices d'un endomorphisme dans deux bases différentes.

5.4.4. Théorème. Soit $T \in \mathrm{End}(E)$. Si $\{u_1, \ldots, u_n\}$ et $\{v_1, \ldots, v_n\}$ sont deux bases de E, alors

$$[T]_{\{v_1,\ldots,v_n\}} = P^{-1}[T]_{\{u_1,\ldots,u_n\}}P,$$

où P est la matrice de passage de $\{u_1, \ldots, u_n\}$ vers $\{v_1, \ldots, v_n\}$.

Démonstration. Posons $A = [T]_{\{u_1,\ldots,u_n\}}$. Soit $P = (P_1 \cdots P_n)$ partagée en colonnes. Alors $v_j = (u_1, \ldots, u_n)P_j$, et donc d'après le théorème 5.4.3, $P^{\{T(v_j)\}}_{\{u_1,\ldots,u_n\}} = AP_j$, $j = 1, \ldots, n$. Ainsi $P^{\{T(v_1),\ldots,T(v_n)\}}_{\{u_1,\ldots,u_n\}} = (AP_1 \cdots AP_n) = AP$. D'où

$$\begin{array}{rcl} (T(v_1), \ldots, T(v_n)) & = & (u_1, \ldots, u_n)(AP) \\ & = & ((v_1, \ldots, v_n)P^{-1})(AP) \\ & = & (v_1, \ldots, v_n)(P^{-1}AP). \end{array}$$

Ainsi $[T]_{\{v_1,\ldots,v_n\}} = P^{-1}AP$. Ceci achève la démonstration.

5.4.5. Définition. On dit que $A, B \in M_n(K)$ sont *semblables*, noté $A \sim B$, s'il existe une matrice inversible P telle que $A = P^{-1}BP$.

5.4.6. Proposition. La relation de semblable de matrices de $M_n(K)$ est une relation d'équivalence.

Démonstration. Soient $A, B, C \in M_n(K)$. D'abord, comme $A = I_n^{-1} A I_n$, on a $A \sim A$. En suite, si $A \sim B$, alors $A = P^{-1}BP$ avec P inversible. Comme $B = PAP^{-1} = Q^{-1}AQ$ avec $Q = P^{-1}$ inversible, on a $B \sim A$. Enfin, si $A \sim B$ et $B \sim C$, alors $A = P^{-1}BP$ et $B = Q^{-1}CQ$. Comme $A = (QP)^{-1}C(QP)$, on a $A \sim C$. Ceci achève la démonstration.

Le Théorème 5.4.4 nous dit que les matrices d'un endomorphisme dans deux bases sont semblables. Réciproquement, on a le résultat suivant.

5.4.7. Proposition. Soit T un endomorphisme de E. Si A est la matrice de T dans une base $\{u_1, \ldots, u_n\}$ de E, alors une matrice $B \in M_n(K)$ est semblable à A si, et seulement si, il existe une base $\{v_1, \ldots, v_n\}$ de E telle que $[T]_{\{v_1,\ldots,v_n\}} = B$.

Démonstration. D'après le théorème 5.4.4, il suffit de montre la necessité. Supposons que $B = P^{-1}AP$ avec P inversible. Posons $(v_1, \ldots, v_n) = (u_1, \ldots, u_n)P$. Alors $\{v_1, \ldots, v_n\}$ est une base de E et P est la matrice de passage de $\{u_1, \ldots, u_n\}$ vers $\{v_1, \ldots, v_n\}$. D'après le théorème 5.4.4, la matrice de T dans $\{v_1, \ldots, v_n\}$ est $P^{-1}AP = B$. Ceci achève la démonstration.

Selon le Théorème 5.4.4 et la proposition 5.4.7, le problème de trouver une base dans laquelle la matrice d'un endomorphisme est le plus simple se ramène au problème de trouver une matrice le plus simple qui est semblable à une matrice carrée. On discutera ce problème dans le chapitre VIII.

5.5. Exercices

1. Soit $M \in M_n(K)$ non nulle. Laquelle(s) parmi les applications $T : M_n(K) \to M_n(K)$ suivantes est linéaire et pourquoi?

   (1) $T(A) = MA$; (2) $T(A) = M + A$.

2. Dans chacun des cas suivants, déterminer si l'application est linéaire ou non:
   (1) $T : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (x + y, y^2)$.
   (2) $\cos : \mathbb{R} \to \mathbb{R} : t \mapsto \cos t$.

3. (1) Vérifier que les vecteurs suivants $u_1 = (-1, 1, 1)$, $u_2 = (1, 2, 1)$, $u_3 = (0, 1, 2)$ forment une base de $\mathbb{R}^3$.
   (2) Trouver une application linéaire $T : \mathbb{R}^3 \to \mathbb{R}^4$ telle que $T(u_1) = T(u_2) = T(u_3) = (1, 0, 2, 1)$.
   (3) Déterminer s'il existe ou non une application linéaire T telle que $T(u_1) = T(u_2) = (1, 2, 0, 1)$, $T(u_3) = (1, 0, 1, 1)$ et $T(-1, 0, 1) = (1, 0, 1, 0)$.

4. Soit $T \in \mathcal{L}(E, F)$, où E et F sont deux K-espaces vectoriels de dimension finie.
   (1) Si T est injective, montrer qu'il existe $S \in \mathcal{L}(F, E)$ tel que $ST = \mathbf{1}_E$.
   (2) Si T est surjective, montrer qu'il existe $S \in \mathcal{L}(F, E)$ tel que $TS = \mathbf{1}_F$.

5. Soient $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^3$ et $S : \mathbb{R}^3 \to \mathbb{R}_2[x]$ des applications linéaires définies par $T_1(a, b) = (2a-b, a, a+b)$, $T_2(a, b) = (a - 3b, b, 2a + 5b)$ et $S(a, b, c) = (a - b + c) + (2a + 3b - c)x$. Calculer $S(6T_1 - 9T_2)$.

6. Soit $E = E_1 \oplus E_2$. Pour tout $u = u_1 + u_2 \in E$ avec $u_i \in E_i,\ i = 1, 2$, définir $p(u) = u_1$. Montrer que p est une application linéaire surjective de E dans E_1.

7. Soient $n \geq m$ des entiers positifs. Considérer $u_i = (a_{i1}, \ldots, a_{im}, \ldots, a_{in}) \in K^n$ et $v_i = (a_{i1}, \ldots, a_{im}) \in K^m$ pour $i = 1, \ldots, r$. Montrer, en construisant une application linéaire de K^n dans K^m, que si $v_1, \ldots, v_r$ sont linéairement indépendants, alors $u_1, \ldots, u_r$ sont également linéairement indépendants.

8. Dans chacun des cas suivants, trouver une base de $\mathrm{Ker}(T)$ et une base de $\mathrm{Im}(T)$, et déterminer si T est injective ou surjective.

   (1) $T : \mathbb{R}^{(4)} \to \mathbb{R}^{(3)} : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 2x_2 + x_4 \\ 2x_1 + 3x_2 + x_3 \\ x_1 + 3x_3 + x_4 \end{pmatrix}.$

   (2) $T : \mathbb{R}^{(4)} \to \mathbb{R}^{(3)} : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 2x_2 + x_4 \\ 2x_1 + 3x_2 + x_3 \\ 3x_1 + 5x_2 + 3x_3 + x_4 \end{pmatrix}.$

9. Considérer l'application linéaire $T : M_2(\mathbb{R}) \to M_2(\mathbb{R}) : A \mapsto AM - MA$, où

$$M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}.$$

   (1) Trouver la matrice de T dans les bases canoniques.
   (2) Trouver une base de $\mathrm{Im}(T)$.
   (3) Trouver une base de $\mathrm{Ker}(T)$.

10. Considérer l'application linéaire suivante:

$$T : \mathbb{R}^4 \to \mathbb{R}^3 : (x_1, x_2, x_3, x_4) \mapsto (x_1 + 2x_2 + 3x_3 + 4x_4, 2x_1 + 3x_2 + 4_3 + 5x_4, 3x_1 + 4x_2 + 5x_3 + 6x_4).$$

   (1) Trouver la matrice de T dans les bases canoniques.
   (2) Trouver une base de $\mathrm{Im}(T)$. Est-ce que T est surjective?
   (3) Trouver une base de $\mathrm{Ker}(T)$. Est-ce que T est injective?

11. Soit $T : \mathbb{R}_4[x] \to \mathbb{R}_3[x]$ une application linéaire dont la matrice dans les bases canoniques est suivante:

$$A = \begin{pmatrix} 1 & 2 & 3 & -1 \\ 3 & 1 & 2 & 1 \\ 4 & 3 & 5 & 0 \end{pmatrix}.$$

   (1) Trouver la formule générale de $T(a + bx + cx^2 + dx^3)$.
   (2) Pour quelle valeur réelle de a, le polynôme $4 + 2x + ax^2$, appartient-il à $\mathrm{Im}(T)$?
   (3) Trouver une base de $\mathrm{Im}(T)$.
   (4) Trouver une base de $\mathrm{Ker}(T)$.

12. Considérer les espaces vectoriels réels $\mathbb{R}_3[x]$ et $\mathbb{R}^{(4)}$.

   (1) Vérifier que $\{f_1 = 1 + 2x + 3x^2, f_2 = 2 + 3x + 4x^2, f_3 = 4 + 5x + 5x^2\}$ est une base de $\mathbb{R}_3[x]$.

(2) Trouver l'application linéaire $T : \mathbb{R}_3[x] \to \mathbb{R}^{(4)}$ dont la matrice dans les bases canoniques est la suivante:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Attention: Il faut spécifier $T(a + bx + cx^2)$.

13. Pour tout $n \geq 1$, construire un isomorphisme $K^n \to K_n[x]$.

14. Considérer l'application linéaire $T : \mathbb{R}_3[x] \to \mathbb{R}^3$ définie par

$$T(a + bx + cx^2) = (a + b - c, a - b + c, -2a + b + c).$$

(1) Vérifier que T est un automorphisme.

(2) Trouver l'inverse T^{-1}. *Attention*: Il faut spécifier $T^{-1}(a, b, c)$.

15. Soit $T : E \to F$ une application linéaire de K-espaces vectoriels de dimension finie.

(1) Supposons que T est surjective. Montrer que F est isomorphe à un complémentaire de $\mathrm{Ker}(T)$.

(2) Supposons que T est injective. Montrer que E est isomorphe à $\mathrm{Im}(T)$.

16. Soit $T \in \mathcal{L}(\mathbb{R}^3,\ \mathbb{R}^{(3)})$ dont la matrice dans les bases canoniques est A. Dans chacun des cas suivants, déterminer si T est un isomorphisme ou non; et si oui, trouver la matrice de son inverse dans les bases canoniques.

$$(1) \quad A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}; \qquad (2) \quad A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix}.$$

17. Soient $T : E \to F$ et $S : F \to G$ des applications linéaires de K-espaces vectoriels de dimension finie.

(1) Si T est surjective, montrer que $\mathrm{rg}(ST) = \mathrm{rg}(S)$.

(2) Si S est injective, montrer que $\mathrm{rg}(ST) = \mathrm{rg}(T)$.

18. Considérer l'endomorphisme T de l'espace vectoriel rationnel $\mathbb{Q}_3[x]$ défini par

$$T(a_0 + a_1x + a_2x^2) = (a_0 + 2a_1 + 3a_2) + (2a_0 - a_1 + a_2)x + (3a_0 + a_1 + 4a_2)x^2.$$

(1) Trouver la matrice de T dans la base canonique $\{1, x, x^2\}$.

(2) Déterminer si T est un automorphisme ou non.

(3) Trouver la matrice de T dans la base $\{1 + x - x^2, 2 - x + x^2, x + 3x^2\}$.

(4) Trouver les coordonnées de $T(2 - 3x - x^2)$ dans $\{1 + x - x^2, 2 - x + x^2, x + 3x^2\}$.

19. (1) Vérifier que $f_1 = 1 - 2x + x^2$, $f_2 = 2 + 3x - x^2$ et $f_3 = 3 + x - 2x^2$ forment une base de l'espace réel $\mathbb{R}_3[x]$.

(2) Trouver l'endomorphisme T de $\mathbb{R}_3[x]$ tel que $T(f_1) = f_1 - f_2$, $T(f_2) = f_2 - f_3$ et $T(f_3) = f_3 - f_1$. *Attention*: It faut donner la formule explicite de $T(a + bx + cx^2)$.

20. Si $A \in M_n(K)$, montrer que l'application $T_A : K^{(n)} \to K^{(n)} : u \mapsto Au$ est un automorphisme de $K^{(n)}$ si et seulement si A est inversible.

21. Soit T un endomorphisme de $\mathbb{Q}^3$ dont la matrice dans la base canonique est

$$\begin{pmatrix} 2 & 0 & -1 \\ 3 & -1 & 4 \\ -4 & 3 & 2 \end{pmatrix}.$$

(1) Vérifier que T est un automorphisme.

(2) Trouver la matrice de T^{-1} dans la base $\{(1,0,1),\ (1,1,0),\ (0,1,1)\}$.

(3) Calculer T^{-2}.

22. Soit T un endomorphisme d'un espace vectoriel E. Montrer que $E = \mathrm{Im}(T) \oplus \mathrm{Ker}(T)$ lorsque $T^2 = T$.

23. Soit T un endomorphisme d'un espace vectoriel complexe E. Si $T^3 = T$, montrer que $E = E_{-1} \oplus E_0 \oplus E_1$, où $E_\lambda = \{u \in E \mid T(u) = \lambda u\}$ pour $\lambda = -1, 0, 1$.

24. Considérer l'espace $K[x]$ des polynômes sur K.

(1) Trouver un endomorphisme de E qui est surjectif et non injectif.

(2) Trouver un endomorphisme de E qui est injectif et non surjectif.

25. Soient A et B deux matrices carrées semblables. Montrer les énoncés suivants.

(1) A et B ont le même rang.

(2) A est inversible si et seulement si B l'est.

(3) A est nilpotente si et seulement si B l'est.

26. Un endomorphisme T de E est dit *nilpotent* si $T^r = \mathbf{0}$ pour un certain $r > 0$. Si E est de dimension finie, montrer que T est nilpotent si et seulement si la matrice de T dans une base de E est nilpotente.

27. Soit $A \in M_2(K)$ non nulle telle que $A^2 = 0$. Montrer que A est semblable à la matrice suivante

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Indication: Considérer l'endomorphisme de $K^{(2)}$ défini par A. Montrer que $Au \neq 0$ pour un certain u et que $\{u, Au\}$ est une base de $K^{(2)}$.

Chapitre VI: Espaces duals

Partout dans ce chapitre, on se fixe K un corps et E un espace vectoriel sur K. Le but de cette section est étudier un nouveau K-espace vectoriel E^* associé à E. Rappelons que K est un espace vectoriel sur lui-même.

6.1. Définition. (1) On appelle $f \in \mathcal{L}(E, K)$ une *forme linéaire* sur E.

(2) Le K-espace vectoriel $E^* = \mathcal{L}(E, K)$, ce qui se compose des formes linéaires sur E, s'appelle *espace dual* de E.

Rappelons que si $\alpha, \beta \in K$ et $f, g \in E^*$, alors $\alpha f + \beta g \in E^*$ est définie par

$$(\alpha f + \beta g)(u) = \alpha f(u) + \beta f(u), \quad \text{pour tout } u \in E.$$

En plus, 0_{E^*} est l'application nulle $0 : E \to K : u \mapsto 0_K$.

Exemple. (1) Soit $M_n(K)$ le K-espace vectoriel des matrices carrées d'ordre n. Alors

$$\mathrm{tr} : M_n(K) \to K : A \mapsto \mathrm{tr}(A)$$

est une forme linéaire sur $M_n(K)$.

(2) Considérons l'espace vectoriel $C[a, b]$ des fonctions continues définies sur l'intervalle $[a, b]$. Alors

$$F : C[a, b] \to \mathbb{R} : f \mapsto \int_a^b f dt$$

est une forme linéaire sur $C[a, b]$. En outre, $2F$ est une forme linéaire telle que $(2F)(f) = 2\int_a^b f dt$.

Le résultat suivant est trivial.

6.2. Lemme. Soit E de dimension finie ayant pour base $\{u_1, \ldots, u_n\}$. Si $f, g \in E^*$, alors $f = g$ si, et seulement si, $f(u_i) = g(u_i)$ pour tout $1 \leq i \leq n$.

Soit $\{u_1, \ldots, u_n\}$ une base de E. Pour tout $1 \leq i \leq n$, posons

$$u_i^* : E \to K : x_1 u_1 + \cdots + x_n u_n \mapsto x_i.$$

Alors $u_i^* \in E^*$ telle que $u_i^*(u_j) = \delta_{ij}$, $j = 1, \ldots, n$, car $u_j = \delta_{1j} u_1 + \cdots + \delta_{nj} u_n$.

6.3. Théorème. Si $\{u_1, \ldots, u_n\}$ est une base de E, alors $\{u_1^*, \ldots, u_n^*\}$ est une base de E^*, appelée la *base duale* de $\{u_1, \ldots, u_n\}$. En particulier, dim(E) =dim(E^*).

Démonstration. Supposons que $\alpha_1 u_1^* + \cdots + \alpha_n u_n^* = 0_{E^*}$, $\alpha_j \in K$. Alors pour $1 \leq i \leq n$,

$$0_K = (\sum_{j=1}^n \alpha_j u_j^*)(u_i) = \sum_{j=1}^n \alpha_j \, u_j^*(u_i) = \sum_{j=1}^n \alpha_j \delta_{ji} = \alpha_i \delta_{ii} = \alpha_i.$$

Donc $\{u_1^*, \ldots, u_n^*\}$ est libre. D'autre part, pour tout $f \in E^*$, posons $f(u_j) = \beta_j \in K, j = 1, \ldots, n$. Alors $\sum_{j=1}^n \beta_j u_j^* \in E^*$ est telle que pour tout $1 \leq i \leq n$,

$$(\sum_{j=1}^n \beta_j u_j^*)(u_i) = \sum_{j=1}^n \beta_j \cdot u_j^*(u_i) = \sum_{j=1}^n \beta_j \cdot \delta_{ji} = \beta_i = f(u_i).$$

Donc $f = \sum_{j=1}^{n} \beta_j u_j^*$. Ce qui montre que $\{u_1^*, \ldots, u_n^*\}$ est une base de E^*.

Exemple. (1) Considérons le K-espace vectoriel K^n. Alors la base duale de la base canonique $\{e_1, \ldots, e_n\}$ est $\{e_1^*, \ldots, e_n^*\}$, où e_i est la fonction suivante:

$$e_i^* : K^n \to K : (x_1, \ldots, x_n) \mapsto x_i,\ i = 1, \ldots, n.$$

(2) Considérons le K-espace vectoriel $K_n[x]$. Alors la base duale de la base canonique $\{1, x, \ldots, x^{n-1}\}$ est $\{1^*, x^*, \cdots, (x^{n-1})^*\}$, où $(x^i)^*$ est la fonction suivante:

$$(x^i)^* : K_n[x] \to K : a_0 + a_1 x + \cdots + a_i x^i + \cdots + a_{n-1} x^{n-1} \to a_i,\ i = 1, \cdots, n.$$

6.4. Proposition. Soient $\{u_1, \ldots, u_n\}$ une base E et $\{u_1^*, \ldots, u_n^*\}$ sa base duale de E^*. Alors une application $f : E \to K$ appartient à E^* si, et seulement si, il existe $a_1, \ldots, a_n \in K$ tels que

$$f(x_1 u_1 + \cdots + x_n u_n) = a_1 x_1 + \cdots + a_n x_n, \text{ pour tous } x_1, \ldots, x_n \in K.$$

Dans ce cas, la colonne des coordonnées de f dans $\{u_1^*, \ldots, u_n^*\}$ est $(a_1, \ldots, a_n)^T$.

Démonstration. Considérant la base $\{u_1, \ldots, u_n\}$ de E et la base $\{1_K\}$ de K, la première partie du lemme suit immédiatement de la proposition 5.1.4. Supposons maintenant que $f : E \to K$ est telle que $f(x_1 u_1 + \cdots + x_n u_n) = a_1 x_1 + \cdots + a_n x_n$, pour tous $x_1, \ldots, x_n \in K$. En particulier, $f(u_i) = a_i = (a_1 u_1^* + \cdots + a_n u_n^*)(u_i)$ pour tout $1 \leq i \leq n$. Donc $f = a_1 u_1^* + \cdots + a_n u_n^*$ comme $\{u_1, \ldots, u_n\}$ est une base de E. Ceci achève la démonstration.

Remarque. Si E est l'un des exemples des espaces vectoriels admettant une base canonique $\{u_1, \ldots, u_n\}$, alors la base duale $\{u_1^*, \ldots, u_n^*\}$ s'appelle *base canonique* de E^*.

Le résultat suivant nous permet de déterminer la base duale d'une base de E.

6.5. Proposition. Soient $\{u_1, \ldots, u_n\}$ et $\{v_1, \ldots, v_n\}$ des bases de E. Si P est la matrice de passage de $\{u_1, \ldots, u_n\}$ vers $\{v_1, \ldots, v_n\}$, alors la matrice de passage de $\{v_1^*, \ldots, v_n^*\}$ vers $\{u_1^*, \ldots, u_n^*\}$ est P^T, et donc la matrice $\{u_1^*, \ldots, u_n^*\}$ vers $\{v_1^*, \ldots, v_n^*\}$ est $(P^T)^{-1}$.

Démonstration. Posons $P = (a_{ij})_{n \times n}$. Soit $Q = (b_{ij})_{n \times n}$ la matrice de passage de $\{v_1^*, \ldots, v_n^*\}$ vers $\{u_1^*, \ldots, u_n^*\}$. Alors $v_j = \sum_{k=1}^{n} a_{kj} u_k$ pour tout $1 \leq j \leq n$ et $u_i^* = \sum_{k=1}^{n} b_{ki} v_k^*$ pour tout $1 \leq i \leq n$. Ainsi pour tous $1 \leq i, j \leq n$,

$$u_i^*(v_j) = \sum\nolimits_{k=1}^{n} a_{kj} u_i^*(u_k) = a_{ij}, \quad u_i^*(v_j) = \sum\nolimits_{k=1}^{n} b_{ki} v_k^*(v_j) = b_{ji},$$

d'où $a_{ij} = b_{ji}$. Ceci montre que $Q = P^T$. La preuve s'achève.

6.6. Corollaire. Si $\{f_1, \ldots, f_n\}$ est une base de E^*, alors il existe une unique base $\{u_1, \ldots, u_n\}$ de E dont la base duale est $\{f_1, \ldots, f_n\}$.

Démonstration. Par hypothèse, $\dim(E) = \dim(E^*) = n$. Prenons une base quelconque $\{v_1, \ldots, v_n\}$ de E. Soit Q la matrice de passage de $\{v_1^*, \ldots, v_n^*\}$ vers $\{f_1, \ldots, f_n\}$. Posons $(u_1, \ldots, u_n) = (v_1, \ldots, v_n)(Q^T)^{-1}$. Alors $\{u_1, \ldots, u_n\}$ est une base de E. Comme $(v_1, \ldots, v_n) = (u_1, \ldots, u_n)Q^T$, la matrice de passage de $\{u_1, \ldots, u_n\}$ vers $\{v_1, \ldots, v_n\}$ est Q^T. D'après la proposition 6.5, la matrice de passage de $\{v_1^*, \ldots, v_n^*\}$ vers $\{u_1^*, \ldots, u_n^*\}$ est $(Q^T)^T = Q$. Ainsi $(u_1^*, \ldots, u_n^*) = (v_1^*, \ldots, v_n^*)Q = (f_1, \ldots, f_n)$. Ainsi $(u_1^*, \ldots, u_n^*) = (f_1, \ldots, f_n)$. Si $\{w_1, \ldots, w_n\}$ est une base de E telle que $(w_1^*, \ldots, w_n^*) = (u_1^*, \ldots, u_n^*)$, alors la matrice de passage $\{u_1^*, \ldots, u_n^*\}$ vers $\{w_1^*, \ldots, w_n^*\}$ est I_n. D'après la proposition 6.5, la matrice de passage de $\{w_1, \ldots, w_n\}$ vers $\{u_1, \ldots, u_n\}$ est $I_n^T = I_n$. Par conséquent, $\{w_1, \ldots, w_n\} = \{u_1, \ldots, u_n\}$. Ceci achève la démonstration.

Soit $T \in \mathcal{L}(E, F)$. Si $f : F \to K$ est une forme linéaire sur F, alors le composé $f\,T : E \to K$ de f et T est une forme linéaire sur E. On définit

$$T^{\,t} : F^* \to E^* : f \mapsto T^{\,t}(f) = f\,T.$$

On a donc, pour tout $f \in F^*$, un diagramme commutatif

$$\begin{tikzcd} E \arrow[r, "T"] \arrow[dr, "T^{\,t}(f)"'] & F \arrow[d, "f"] \\ & K. \end{tikzcd}$$

6.7. Proposition. Si $T \in \mathcal{L}(E, F)$, alors

$$T^{\,t} : F^* \to E^* : f \mapsto T^{\,t}(f)$$

est une application linéaire, appelée la *transposée* de T.

Démonstration. Si $f, g \in F^*, \alpha, \beta \in K$, alors $T^{\,t}(\alpha f + \beta g),\ \alpha T^{\,t}(f) + \beta T^{\,t}(g) \in E^*$. Pour tout $u \in E$,

$$\begin{aligned} T^{\,t}(\alpha f + \beta g)(u) &= ((\alpha f + \beta g)\,T)\,(u) = (\alpha f + \beta g)(T(u)) \\ &= \alpha f(T(u)) + \beta g(T(u)) = \alpha(f\,T)(u) + \beta(g\,T)(u) \\ &= \alpha T^{\,t}(f)(u) + \beta T^{\,t}(g)(u) = (\alpha T^{\,t}(f) + \beta T^{\,t}(g))\,(u). \end{aligned}$$

Donc $T^{\,t}(\alpha f + \beta g) = \alpha T^{\,t}(f) + \beta T^{\,t}(g)$, c'est-à-dire, $T^{\,t}$ est linéaire.

6.8. Théorème. Soient $\{u_1, \ldots, u_n\}$ une base de E et $\{v_1, \ldots, v_m\}$ une base de F. Si $T \in \mathcal{L}(E, F)$ dont la matrice dans les bases $\{u_1, \ldots, u_n\}$ et $\{v_1, \ldots, v_m\}$ est A, alors la matrice de $T^{\,t}$ dans les bases duales $\{v_1^*, \ldots, v_m^*\}$ et $\{u_1^*, \ldots, u_n^*\}$ est A^T.

Démonstration. Posons $A = (a_{ij})_{m\times n}$. Alors $T(u_i) = \sum_{k=1}^m a_{ki} v_k,\ i = 1, \ldots, n$. Soit $B = (b_{ij})_{n\times m}$ la matrice de $T^{\,t}$ dans les bases $\{v_1^*, \ldots, v_m^*\}$ et $\{u_1^*, \ldots, u_n^*\}$. Alors $T^{\,t}(v_j^*) = \sum_{k=1}^n b_{kj} u_k^*,\ j = 1, \ldots, m$. On se fixe $1 \le i \le n,\ 1 \le j \le m$. Alors

$$T^{\,t}(v_j^*)(u_i) = \left(\sum_{k=1}^n b_{kj} u_k^*\right)(u_i) = \sum\nolimits_{k=1}^n b_{kj} u_k^*(u_i) = \sum\nolimits_{k=1}^n b_{kj}\delta_{ki} = b_{ij} 1_K = b_{ij}.$$

D'autre part,

$$T^{\,t}(v_j^*)(u_i) = v_j^*\,(T(u_i)) = v_j^*\left(\sum\nolimits_{k=1}^m a_{ki} v_k\right) = \sum\nolimits_{k=1}^m a_{ki} v_j^*(v_k) = \sum\nolimits_{k=1}^m a_{ki}\delta_{jk} = a_{ji} 1_K = a_{ji}.$$

D'où $b_{ij} = a_{ji}$. Ce qui achève la démonstration.

6.9. Corollaire. Soient E et F des K-espaces vectoriels de dimension finie. Pour toute $S \in \mathcal{L}(F^*, E^*)$, il existe une unique $T \in \mathcal{L}(E, F)$ telle que $T^t = S$.

Démonstratio. Prenons une base $\{u_1, \ldots, u_n\}$ de E et une base $\{v_1, \ldots, v_m\}$ de F. Supposons $S \in \mathcal{L}(F^*, E^*)$ dont la matrice dans les bases $\{v_1^*, \ldots, v_m^*\}$ et $\{u_1^*, \ldots, u_n^*\}$ est A. A'après le lemme 5.3.9, il existe une unique $T \in \mathcal{L}(E, F)$ dont la matrice dans les bases $\{u_1, \ldots, u_n\}$ et $\{v_1, \ldots, v_m\}$ est A^T. Or la matrice de T^t dans les bases $\{v_1^*, \ldots, v_m^*\}$ et $\{u_1^*, \ldots, u_n^*\}$ est $(A^T)^T = A$. D'où $T^t = S$. Ceci achève la démonstration.

Remarquons que si $T \in \mathrm{End}(E)$, alors $T^{\,t} \in \mathrm{End}(E^*)$. Le résultat suivant découle immédiatement du théorème 6.8 et le corollaire 6.9.

6.10. Théorème. Soit E de dimension finie ayant pour base $\{u_1, \ldots, u_n\}$.

(1) Si $T \in \mathrm{End}(E)$ dont la matrice dans $\{u_1, \ldots, u_n\}$ est A, alors la matrice de $T^{\,t} \in \mathrm{End}(E^*)$ dans la base duale $\{u_1^*, \ldots, u_n^*\}$ est A^T.

(2) Pour tout $S \in \mathrm{End}(E^*)$, il existe un unique $T \in \mathrm{End}(E)$ tel que $T^t = S$.

6.11. Exercices

1. Soient f et g les formes linéaires sur $\mathbb{R}^2$ définies par $f(x, y) = 2x - 3y$ et $g(x, y) = 3x + y$.
   (1) Calculer $2f + 3g$.
   (2) Vérifier que f et g sont linéairement indépendants.

2. Montrer premièrement que $\{(1, -2, 3), (1, -1, 1), (2, -4, 7)\}$ est une base de $\mathbb{Q}^3$, et ensuite trouver sa base duale.

3. Considérer les formes linéaires f_1, f_2 et f_3 définies, pour tout $u = (x_1, x_2, x_3) \in \mathbb{R}^3$, par
$$f_1(u) = x_1 + 2x_2 + x_3,\ f_2(u) = 2x_1 + 3x_2 + 3x_3,\ f_3(u) = 3x_1 + 7x_2 + x_3.$$
   Vérifier que $\{f_1, f_2, f_3\}$ est une base de l'espace dual de $\mathbb{R}^3$ et trouver la base de $\mathbb{R}^3$ dont la base duale est $\{f_1, f_2, f_3\}$.

4. Soient f_1, f_2, g les formes linéaires sur l'espace réel $\mathbb{C}$ définies par $f_1(a+bi) = 3a-4b,\ f_2(a+bi) = 4a+2b$ et $g(a + bi) = 3a - 7b$.
   (1) Vérifier que $\{f_1, f_2\}$ est une base de l'espace dual de $\mathbb{C}$.
   (2) Trouver les coordonnées de g dans cette base.

5. Soit E un K-espace vectoriel de dimension finie. Si $u \in E$, montrer que $u = 0_E$ si et seulement si $f(u) = 0_K$ pour tout $f \in E^*$.

6. Soit E un espace vectoriel de dimension n, et soient $f_1, \ldots, f_n \in E^*$.
   (1) Si $f_1, \ldots, f_n$ s'annulent sur un même vecteur non nul de E, montrer qu'elles sont linéairement dépendantes. *Indication*: Utiliser la question précédente.
   (2) Si $u_1, \ldots, u_n \in E$ sont tels que $f_i(u_j) = \delta_{ij}$ pour tous $1 \le i, j \le n$, montrer que $\{u_1, \ldots, u_n\}$ est une base de E et $\{f_1, \ldots, f_n\}$ est une base de E^*.

7. Considérer l'espace réel $\mathbb{R}^3 = \{(x_1, x_2, x_3) \mid x_i \in \mathbb{R}\}$ et son espace dual $(\mathbb{R}^3)^*$. Soient $g_1, g_2, g_3 \in (\mathbb{R}^3)^*$ définies par
$$g_1(x_1, x_2, x_3) = 3x_1 - 2x_2,\ g_2(x_1, x_2, x_3) = 2x_2 - x_3,\ g_3(x_1, x_2, x_3) = x_1 + x_3.$$
   (1) Vérifier que $\{g_1, g_2, g_3\}$ est une base de $(\mathbb{R}^3)^*$.

(2) Trouver la matrice de passage de $\{g_1, g_2, g_3\}$ à la base duale $\{e_1^*, e_2^*, e_3^*\}$ de la base canonique de $\mathbb{R}^3$.

(3) Soit $T \in \mathcal{L}(\mathbb{R}^3,\ \mathbb{R}^3)$ défini par $T(x_1, x_2, x_3) = (-3x_2 + x_3, 2x_1 + x_3, x_1 - x_2)$. Trouver la matrice de la transposée $T^{\,t}$ dans la base $\{g_1, g_2, g_3\}$.

(4) Soit G l'endomorphisme de $(\mathbb{R}^3)^*$ défini par

$$G(ag_1 + bg_2 + cg_3) = (a - c)g_1 + (b - a)g_2 + (c - b)g_3.$$

Déterminer l'endomorphisme S de $\mathbb{R}^3$ dont la transposée est G.

8. Soient E et F des K-espaces vectoriels de dimension finie. Si $T \in \mathcal{L}(E, F)$, montrer que T est un isomorphisme si et seulement si sa transposée $T^t : F^* \to E^*$ est un isomorphisme.

9. Soit E un espace de dimension finie. Si F est un sous-espace de E, montrer que

$$F^\perp = \{f \in E^* \mid f(u) = 0,\ \text{pour tout } u \in F\}$$

est un sous-espace de E^* et que $\dim(E) = \dim(F^\perp) + \dim(F)$.

Indication: Si $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ une base de E avec $\{u_1, \ldots, u_r\}$ une base de F, vérifier que $\{u_{r+1}^*, \ldots, u_n^*\}$ est une base de $F^\perp$.

10. Soit F le sous-espace vectoriel de l'espace réel $\mathbb{R}_4[x]$ engendré par $u_1 = 1+x-x^2+2x^3$, $u_2 = 2x+3x^2-x^3$, $u_3 = 1 + 3x + 2x^2 + x^3$, $u_4 = 1 + 5x + 5x^2$. Trouver une base de $F^\perp$.

Chapitre VII: Déterminants

Les déterminants de matrices réelles d'ordre 2 et 3 sont apparues en relation avec le calcul d'aire et volume, respectivement. Dans ce chapitre, nous nous proposons de définir les déterminants de matrices carrées sur un corps quelconque d'ordre n pour tout $n \geq 1$. Nous étudierons les développements et appliquerons les résultats obtenus à la résolution de certain systèmes d'équations linéaires, ainsi qu'au calcul de l'inverse d'une matrice inversible.

Partout dans ce chapitre, K désignera un corps.

7.1. Permutations

On définira les déterminants d'ordre n au moyen de permutations sur n lettres. Dans cette section, on rappelle rapidement la notion de permutations et certains propriétés que l'on aura besoin pour étudier les déterminants.

7.1.1. Définition. Soit $n \geq 1$. Une *n-permutation* est une application bijective

$$\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\} : i \mapsto \sigma(i).$$

Dans ce cas, on note

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

7.1.2. Proposition. L'ensemble S_n des n-permutations est un groupe d'ordre $n!$ pour la composition d'applications.

7.1.3. Définition. On appelle $\sigma \in S_n$ un *cycle de longueur* $r \geq 2$ s'il existe $i_1, i_2, \ldots, i_r$ distincts de $\{1, 2, \ldots, n\}$ tels que $\sigma(i_1) = i_2, \ldots, \sigma(i_{r-1}) = i_r,\ \sigma(i_r) = i_1$ et $\sigma(i) = i$, pour tout $i \notin \{i_1, i_2, \ldots, i_r\}$. Dans ce cas, on note $\sigma = (i_1 i_2 \cdots i_r)$. En outre on appelle un cycle (ij) de longueur 2 *transposition.*

Remarque. (1) Pour tout $i \in \{1, 2, \ldots, n\}$, on définit (i), un cycle de longueur 1, comme étant l'application identité:

$$(i) = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

(2) Pour tous $i, j \in \{1, \cdots, n\}$, on a $(ij) = (ji)$ et $(ij)^{-1} = (ij)$.

7.1.4. Lemme. Toute permutation se décompose en produit de cycles.

Démonstration. Soit $\sigma \in S_n$. Soit $r(\sigma)$ le nombre des entiers $i \in \{1, \ldots, n\}$ tel que $\sigma(i) \neq i$. On procède par récurrence sur $r(\sigma)$. Si $r(\sigma) = 0$, alors $\sigma = (1)$. Supposons que $r > 0$ et la proposition est vraie pour $\tau \in S_n$ avec $r(\tau) < r$. Soit $r(\sigma) = r$. Alors il existe un $1 \leq a \leq n$ tel que $\sigma(a) \neq a$. Or il existe un $s > 1$ tel que $\sigma^s(a) = a$ et $\sigma^{s-1}(a) \neq a$. Posons $a_j = \sigma^{j-1}(a)$ pour tout $1 \leq j \leq s$. Alors $a_1, a_2, \ldots, a_s$ sont deux à deux distincts et $\sigma(a_j) = a_{j+1}$ pour $1 \leq j < s$ et $\sigma(a_s) = a_1$. Définissons $\tau \in S_n$ par $\tau(i) = \sigma(i)$ si $i \notin \{a_1, \ldots, a_s\}$ et $\tau(i) = i$ si $i \in \{a_1, \ldots, a_s\}$. Il est facile de vérifier que $\sigma = (a_1 a_2 \cdots a_s)\tau$. Remarquons que $r(\tau) < r(\sigma) = r$. Ainsi τ se decompose en produit de cycles. Il en est de même pour σ. Ce qui achève la démonstration.

La démonstration du lemme suivant est une vérification de routine.

7.1.5. Lemme. Un cycle $(i_1 i_2 \cdots i_r)$ avec $r \geq 2$ se décompose

$$(i_1 i_2 \cdots i_r) = (i_1 i_r)(i_1 i_{r-1}) \cdots (i_1 i_2).$$

On accepte le résultat sans démonstration.

7.1.6. Théorème. Si $n \geq 2$, alors toute n-permutation se décompose en un produit de transpositions dont la parité du nombre de transpositions est invariante.

7.1.7. Définition. Si $n \geq 2$, alors une n-permutation est dite *paire* si elle se décompose en un produit d'un nombre pair de transpositions; et *impaire* sinon. Par convention, la seule 1-permutation (1) est dite paire.

Remarque. Si $n \geq 2$, alors la n-permutation identité (1) est paire puisque $(1) = (12)(12)$.

7.1.8. Définition. On définit la *signature* d'une permutation σ par

$$\operatorname{sgn}(\sigma) = \begin{cases} +1, & \text{si } \sigma \text{ est paire} \\ -1, & \text{si } \sigma \text{ est impaire.} \end{cases}$$

Remarque. Si $\sigma = (a_1 b_1) \cdots (a_r b_r)$, alors $\operatorname{sgn}(\sigma) = (-1)^r$.

7.1.9. Proposition. Si $\sigma, \tau \in S_n$, alors $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$ et $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau)$.
Démonstration. Supposons que $\sigma = (a_1 b_1) \cdots (a_r b_r)$ et $\tau = (c_1 d_1) \cdots (c_s d_s)$. Alors $\sigma^{-1} = (a_r b_r) \cdots (a_1 b_1)$ et $\sigma\tau = (a_1 b_1) \cdots (a_r b_r)(c_1 d_1) \cdots (c_s d_s)$. Par conséquent, on voit que $\operatorname{sgn}(\sigma^{-1}) = (-1)^r = \operatorname{sgn}(\sigma)$, et

$$\operatorname{sgn}(\sigma\tau) = (-1)^{r+s} = (-1)^r(-1)^s = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau).$$

Ceci achève la démonstration.

7.2. Déterminant d'une matrice carrée

7.2.1. Définition. Soit $A = (a_{ij})_{n\times n} \in M_n(K)$. Le *déterminant* de A, noté $\det(A)$ ou $|a_{ij}|_{n\times n}$, est défini par

$$\det(A) = \textstyle\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \in K.$$

Remarque. La définition ci-dessus nous donne une fonction

$$\det : M_n(K) \to K : A \mapsto \det(A).$$

Exemple. (1) Comme $S_1 = \{e = (1)\}$, on a $\det(a_{11}) = \operatorname{sgn}(e) a_{1e(1)} = (+1) a_{11} = a_{11}$.

(2) Comme $S_2 = \{e = (1), \sigma = (12)\}$, on a

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \mathrm{sgn}(e) a_{1e(1)} a_{2e(2)} + \mathrm{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} = a_{11}a_{22} - a_{12}a_{21}.$$

(3) Comme $S_3 = \{\sigma_1 = (1), \sigma_2 = (12), \sigma_3 = (13), \sigma_4 = (23), \sigma_5 = (123), \sigma_6 = (132)\}$, pour toute $A = (a_{ij})_{3\times 3}$, on a

$$\begin{array}{rcl} \det(A) & = & \sum_{i=1}^{6} \mathrm{sgn}(\sigma_i)\, a_{1\sigma_i(1)} a_{2\sigma_i(2)} a_{3\sigma_i(3)} \\ & = & a_{11}a_{22}a_{33} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & = & (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{12}a_{21}a_{33} + a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32}). \end{array}$$

Ceci est la différence de la somme de trois produits des terms diagonaux et la somme de trois produits des termes anti-diagonaux de la matrice suivante:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{pmatrix}.$$

7.2.2. Théorème. Pour toute $A = (a_{ij}) \in M_n(K)$, on a

$$\det(A) = \sum\nolimits_{\sigma \in S_n} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Démonstration. Pour $\sigma \in S_n$, posons $j_i = \sigma(i)$, c'est-à-dire, $i = \sigma^{-1}(j_i), i = 1, 2, \ldots, n$. De cette manière, on a

$$a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = a_{j_1, \sigma^{-1}(j_1)} a_{j_2, \sigma^{-1}(j_2)} \cdots a_{j_n, \sigma^{-1}(j_n)} = a_{1, \sigma^{-1}(1)} a_{2, \sigma^{-1}(2)} \cdots a_{n, \sigma^{-1}(n)},$$

où la dernière égalité suit du fait que $\{j_1, j_2, \ldots, j_n\} = \{1, 2, \ldots, n\}$. Quand σ parcourt S_n, on voit que σ^{-1} parcourt S_n. Donc

$$\begin{array}{rcl} \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} & = & \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\, a_{1, \sigma^{-1}(1)} a_{2, \sigma^{-1}(2)} \cdots a_{n, \sigma^{-1}(n)} \\ & \overset{\tau = \sigma^{-1}}{=} & \sum_{\tau \in S_n} \mathrm{sgn}(\tau)\, a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)} \\ & = & \det(A). \end{array}$$

Ceci achève la démonstration.

Remarque. Si $n = 3$, on peut vérifier directement le théorème 7.2.2 comme suit:

$$\begin{array}{rl} & \sum_{i=1}^{6} \mathrm{sgn}(\sigma_i)\, a_{\sigma_i(1)1} a_{\sigma_i(2)2} a_{\sigma_i(3)3} \\ = & a_{11}a_{22}a_{33} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ = & |a_{ij}|_{3\times 3}. \end{array}$$

7.2.3. Corollaire. Pour toute $A \in M_n(K)$, on a $\det(A) = \det(A^T)$.

Démonstration. Écrivons $A = (a_{ij})_{n\times n}$. Alors $A^T = (b_{ij})$, où $b_{ij} = a_{ji}$, pour tous $1 \le i, j \le n$. Or

$$\begin{array}{rcl} \det(A^T) & = & \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\, b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} \\ & = & \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \\ & = & \det(A). \end{array}$$

Ceci achève la démonstration.

On étudiera le calcul du déterminant.

7.2.4. Proposition. Si $A = (a_{ij}) \in M_n(K)$ est triangulaire, alors $\det(A) = a_{11}a_{22}\cdots a_{nn}$.

Démonstration. D'abord, supposons que A est triangulaire inférieure, c'est-à-dire, $a_{ij} = 0$, pour $1 \le i < j \le n$. Soit $\sigma \in S_n$. Si $\sigma \ne (1)$, alors il existe un élément minimal r de $\{1, \ldots, n\}$ tel que $\sigma(r) \ne r$. D'après la minimalité, $\sigma(i) = i$ pour tout $1 \le i < r$. Ainsi $\sigma(r) \notin \{1, \ldots, r-1, r\}$. D'où, $\sigma(r) > r$, et donc $a_{r\sigma(r)} = 0$. Par conséquent, $a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} = 0$. Ceci nous donne

$$\begin{array}{rcl}\det(A) & = & \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} \\ & = & a_{11}a_{22}\cdots a_{nn} + \sum_{(1)\ne\sigma\in S_n,} \mathrm{sgn}(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} \\ & = & a_{11}a_{22}\cdots a_{nn}.\end{array}$$

Si A est triangulaire supérieure, alors $A^T = (a_{ji})$ est triangulaire inférieure. D'après le corollaire 7.2.3, on a $\det(A) = \det(A^T) = a_{11}\cdots a_{nn}$. Ceci achève la démonstration.

Exemple. $\det(I_n) = 1_K$.

7.2.5. Lemme. Si l'on échange deux colonnes ou deux lignes d'une matrice carrée, alors le déterminant change de signe.

Démonstration. Soit $B = (b_{ij})$ obtenue à partir de $A = (a_{ij})$ en échangeant la r-ième colonne et la s-ième colonne avec $1 \le r < s \le n$. Alors $b_{ir} = a_{is}, b_{is} = a_{ir}$, et $b_{ij} = a_{ij} =$ lorsque $j \ne r, s$. Comme S_n est un groupe, on voit que $\sigma(rs)$ parcourt S_n lorsque σ parcourt S_n. Donc

$$\begin{array}{rcl}\det(B) & = & \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma)\, b_{\sigma(1)1}\cdots b_{\sigma(r)r}\cdots b_{\sigma(s)s}\cdots b_{\sigma(n)n} \\ & \overset{\sigma(rs)=\tau}{=} & -\sum_{\tau\in S_n} \mathrm{sgn}(\tau)\, b_{\tau(1)1}\cdots b_{\tau(s)r}\cdots b_{\tau(r)s}\cdots b_{\tau(n)n} \\ & = & -\sum_{\tau\in S_n} \mathrm{sgn}(\tau)\, a_{\tau(1)1}\cdots a_{\tau(s)s}\cdots a_{\tau(r)r}\cdots a_{\tau(n)n} \\ & = & -\sum_{\tau\in S_n} \mathrm{sgn}(\tau)\, a_{\tau(1)1}\cdots a_{\tau(r)r}\cdots a_{\tau(s)s}\cdots a_{\tau(n)n} \\ & = & -\det(A).\end{array}$$

Ceci achève la démonstration.

7.2.6. Lemme. Soit $A \in M_n(K)$. Si deux colonnes ou deux lignes de A sont identiques, alors $\det(A)$ est nul.

Démonstration. Soit $A = (a_{ij})_{n\times n}$. Supposons qu'il existe r, s avec $1 \le r < s \le n$ tels que $a_{ir} = a_{is}$, pour tout $1 \le i \le n$. Soient $\sigma_1, \ldots, \sigma_m$ les permutations paires de S_n. Posant $\tau_j = \sigma_j(rs)$, on voit que $\tau_1, \ldots, \tau_m$ sont les permutations impaires de S_n. Pour tout $1 \le i \le m$, on a

$$\begin{array}{rl} & \mathrm{sgn}(\tau_i)a_{\tau_i(1)1}\cdots a_{\tau_i(r)r}\cdots a_{\tau_i(s)s}\cdots a_{\tau_i(n)n} \\ = & -\mathrm{sgn}(\sigma_i)a_{\sigma_i(1)1}\cdots a_{\sigma_i(s)r}\cdots a_{\sigma_i(r)s}\cdots a_{\sigma_i(n)n} \\ = & -\mathrm{sgn}(\sigma_i)a_{\sigma_i(1)1}\cdots a_{\sigma_i(s)s}\cdots a_{\sigma_i(r)r}\cdots a_{\sigma_i(n)n} \\ = & -\mathrm{sgn}(\sigma_i)a_{\sigma_i(1)1}\cdots a_{\sigma_i(r)r}\cdots a_{\sigma_i(s)s}\cdots a_{\sigma_i(n)n}.\end{array}$$

Comme $S_n = \{\sigma_j, \tau_j \mid j = 1, \ldots, m\}$, on a

$$\det(A) = \sum_{i=1}^{m} \left(\mathrm{sgn}(\sigma_i)a_{\sigma_i(1)1}\cdots a_{\sigma_i(n)n} + \mathrm{sgn}(\tau_i)a_{\tau_i(1)1}\cdots a_{\tau_i(n)n}\right) = 0.$$

Ceci achève la démonstration.

7.2.7. Lemme. Soit $A = (A_1, \dots, A_n) \in M_n(K)$ partagée en colonnes. Si $A_s = aB_s + bC_s$ pour un certain $1 \le s \le n$, où $a, b \in K$ et B_s, C_s des matrices-colonnes, alors

$$\det(A) = a\det(A_1, \dots, A_{s-1}, B_s, A_{s+1}, \dots, A_n) + b\det(A_1, \dots, A_{s-1}, C_s, A_{s+1}, \dots, A_n).$$

L'énoncé correspondant pour les lignes est également valide.

Démonstration. Posons $A = (a_{ij})_{n\times n}$, $B = (A_1, \dots, A_{s-1}, B_s, A_{s+1}, \dots, A_n) = (b_{ij})_{n\times n}$ et $C = (A_1, \dots, A_{s-1}, C_s, A_{s+1}, \dots, A_n) = (c_{ij})_{n\times n}$. Alors, pour tout $1 \le i \le n$, on a $a_{is} = ab_{is} + bc_{is}$, et $a_{ij} = b_{ij} = c_{ij}$ lorsque $j \ne s$. Or

$$\begin{array}{rcl} \det(A) & = & \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(s)s} \cdots a_{\sigma(n)n} \\ & = & \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots (ab_{\sigma(s)s} + bc_{\sigma(s)s}) \cdots a_{\sigma(n)n} \\ & = & a\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(s),s} \cdots b_{\sigma(n)n} + b\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) c_{\sigma(1)1} \cdots c_{\sigma(s),s} \cdots c_{\sigma(n)n} \\ & = & a\det(B) + b\det(C). \end{array}$$

Ceci achève la démonstration.

Remarque. Le lemme 7.2.7 dit que la fonction $\det : M_n(K) \to K : A \mapsto \det(A)$ est linéaire par rapport à chacun des lignes et chacun des colonnes.

7.2.8. Corollaire. (1) $\det(A_1, \cdots, aA_s, \cdots, A_n) = a\det(A_1, \cdots, A_s, \cdots, A_n)$. Il en est de même pour les lignes de A.

(2) Si une colonne ou une ligne de A est nulle, alors $\det(A) = 0$.

(3) Pour toute $A \in M_n(K)$ et tout $a \in K$, $\det(aA) = a^n \det(A)$.

7.2.9. Théorème. Si l'on additionne à une colonne (respectivement, à une ligne) d'une matrice carrée une combinaison linéaire des autres colonnes (respectivement, des autres lignes), alors le déterminant ne change pas.

Démonstration. Soit $A = (A_1, \dots, A_n) \in M_n(K)$ partagée en colonnes. Posons $B = (A_1 \cdots A_{i-1}, A_i + \sum_{j\ne i} a_j A_j, A_{i+1}, \cdots, A_n)$. Il suit des lemmes 7.2.7 et 7.2.6 que

$$\begin{array}{rcl} \det(B) & = & \det(A_1 \cdots A_{i-1}, A_i, A_{i+1}, \cdots, A_n) + a_j \sum_{j\ne i} \det(A_1 \cdots A_{i-1}, A_j, A_{i+1}, \cdots, A_n) \\ & = & \det(A). \end{array}$$

Ceci achève la démonstration.

Exemple. Soit $n \ge 2$. Montrer que

$$V_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{vmatrix} = \Pi_{1\le i<j\le n}(\alpha_j - \alpha_i).$$

Ceci s'appelle le *déterminant de Vandemonde.*

Démonstration. D'abord, il est évident que $V_2 = (\alpha_2 - \alpha_1)$. Supposons que $n > 2$ et $V_{n-1} = \Pi_{1\le i<j\le n-1}(\alpha_j - \alpha_i)$. En effectuant successivement les opérations $L_n - \alpha_1 L_{n-1},\ L_{n-1} - \alpha_1 L_{n-2}, \dots, L_2 - \alpha_1 L_1$, on obtient

$$V_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & \alpha_2 - \alpha_1 & \cdots & \alpha_n - \alpha_1 \\ 0 & \alpha_2^2 - \alpha_1\alpha_2 & \cdots & \alpha_n^2 - \alpha_1\alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_2^{n-1} - \alpha_1\alpha_2^{n-2} & \cdots & \alpha_n^{n-1} - \alpha_1\alpha_n^{n-2} \end{vmatrix}$$

$$= \begin{vmatrix} \alpha_2 - \alpha_1 & \alpha_3 - \alpha_1 & \cdots & \alpha_n - \alpha_1 \\ \alpha_2(\alpha_2 - \alpha_1) & \alpha_3(\alpha_3 - \alpha_1) & \cdots & \alpha_n(\alpha_n - \alpha_1) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_2^{n-2}(\alpha_2 - \alpha_1) & \alpha_3^{n-2}(\alpha_3 - \alpha_1) & \cdots & \alpha_n^{n-2}(\alpha_n - \alpha_1) \end{vmatrix}$$

$$= \Pi_{j=2}^n(\alpha_j - \alpha_1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_2^{n-2} & \alpha_3^{n-2} & \cdots & \alpha_n^{n-2} \end{vmatrix}$$

$$= \Pi_{j=2}^n(\alpha_j - \alpha_1)\, \Pi_{2\le i<j\le n}(\alpha_j - \alpha_i) = \Pi_{1\le i<j\le n}(\alpha_j - \alpha_i).$$

Pour conclure cette section, on montrera que le déterminant du produit de matrices est égal au produit de déterminants.

7.2.10. Définition. Soit $A = (A_1, \cdots, A_n) \in M_n(K)$ partagée en colonne. Pour $\sigma \in S_n$, on définit $\sigma \cdot A = (A_{\sigma(1)} \cdots A_{\sigma(n)}) \in M_n(K)$.

7.2.11. Lemme. Si $A \in M_n(K)$, alors $\det(\sigma \cdot A) = \operatorname{sgn}(\sigma)\det(A)$, pour tout $\sigma \in S_n$.
Démonstration. Partageons $A = (A_1, \cdots, A_n)$ en colonnes. Écrivons $\sigma = (r_1\, s_1) \cdots (r_p\, s_p)$ avec $r_i < s_i$, pour tout $1 \le i \le p$. Si $p = 1$, alors

$$\begin{array}{rcl} \det(\sigma \cdot A) & = & \det(A_1, \cdots, A_{s_1}, \cdots, A_{r_1}, \cdots, A_n) \\ & = & -\ \det(A_1, \cdots, A_{r_1}, \cdots, A_{s_1}, \cdots, A_n) \\ & = & -\ \det(A) = \operatorname{sgn}(\sigma)\det(A). \end{array}$$

Supposons que $p > 1$ et le résultat est vrai pour $p - 1$. Posant $\tau = (r_1 s_1) \cdots (r_{p-1} s_{p-1})$, on a $\sigma = \tau(r_p s_p)$ et

$$\begin{array}{rcl} \det(\sigma \cdot A) & = & \det(A_{\tau(1)} \cdots A_{\tau(s_p)} \cdots A_{\tau(r_p)} \cdots A_{\tau(n)}) \\ & = & -\det(A_{\tau(1)} \cdots A_{\tau(r_p)} \cdots A_{\tau(s_p)} \cdots A_{\tau(n)}) \\ & = & -\det(\tau \cdot A) = -\operatorname{sgn}(\tau)\det(A) = \operatorname{sgn}(\sigma)\det(A). \end{array}$$

Ceci achève la démonstration.

7.2.12. Théorème. Si A, $B \in M_n(K)$, alors $\det(AB) = \det(A)\det(B)$.
Démonstration. Partagóns $A = (A_1, \ldots, A_n)$ en colonnes et écrivons $B = (b_{ij})_{n\times n}$. Alors

$$AB = (A_1, \ldots, A_n) \cdot (b_{ij})_{n\times n} = \left(\sum\nolimits_{i_1=1}^n A_{i_1} b_{i_1 1}, \cdots, \sum\nolimits_{i_n=1}^n A_{i_n} b_{i_n n}\right).$$

Il suit du lemme 7.2.7 que

$$\det(AB) = \sum\nolimits_{i_1=1}^{n} \cdots \sum\nolimits_{i_n=1}^{n} b_{i_1,1} \cdots b_{i_n,n} \det(A_{i_1}, \cdots, A_{i_n}).$$

Si $i_1, \ldots, i_n$ ne sont pas distincts, alors $\det(A_{i_1}, \ldots, A_{i_n}) = 0_K$. De l'autre côté, $i_1, \ldots, i_n$ sont distincts si, et seulement si, il existe $\sigma \in S_n$ tel que $i_j = \sigma(j)$, $j = 1, \ldots, n$. Ainsi

$$\begin{array}{rcl} \det(AB) & = & \sum_{\sigma \in S_n} b_{\sigma(1),1} \cdots b_{\sigma(n),n} \det(A_{\sigma(1)}, \ldots, A_{\sigma(n)}) \\ & = & \sum_{\sigma \in S_n} b_{\sigma(1),1} \cdots b_{\sigma(n),n} \det(\sigma \cdot A) \\ & = & \sum_{\sigma \in S_n} b_{\sigma(1),1} \cdots b_{\sigma(n),n} \operatorname{sgn}(\sigma) \det(A) \\ & = & \det(A) \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n} \right) = \det(A)\det(B). \end{array}$$

Ceci achève la démonstration.

7.3. Expansion de Laplace

Cette section a pour but d'étudier le calcul de déterminants par induction. Soit A une matrice sur K. Une matrice B s'appelle une *sous-matrice* de A si B est obtenue à partir de A en supprimant des lignes et des colonnes. Un *mineur* d'ordre r de A est le déterminant d'une sous-matrice carrée d'ordre r de A.

7.3.1. Définition. Soit $A \in M_n(K)$ avec $n \geq 2$. Pour tous $1 \leq i, j \leq n$, on appelle (i, j)-*mineur* de A, noté $m_{ij}(A)$, le déterminant de la sous-matrice carrée d'ordre $n-1$ obtenue à partir de A en supprimant la i-ième ligne et la j-ième colonne. En outre, $c_{ij}(A) = (-1)^{i+j} m_{ij}(A)$ s'appelle (i, j)-*cofacteur* de A.

Remarque. Soient $A, B \in M_n(K)$. Si A, B ont les mêmes lignes sauf la r-ième, alors $m_{rj}(A) = m_{rj}(B)$ et $c_{rj}(A) = c_{rj}(B)$, pour tout $1 \leq j \leq n$. De même, si A, B ont les mêmes colonnes sauf la s-ième, alors $m_{is}(A) = m_{is}(B)$ et $c_{is}(A) = c_{is}(B)$, pour tout $1 \leq i \leq n$.

7.3.2. Lemme. Soit $A = (a_{ij}) \in M_n(K)$. Si la n-ième ligne de A est de la forme $(0, \ldots, 0, a_{nn})$, alors $\det(A) = a_{nn} m_{nn}(A)$.

Démonstration. Posons $M = (a_{ij})_{(n-1)\times(n-1)}$. Alors $m_{nn}(A) = \det(M)$. Posons $G = \{\sigma \in S_n \mid \sigma(n) = n\}$. On peut identifier G avec S_{n-1}. Si $\sigma \in S_n \backslash G$, alors $\sigma(n) < n$, et donc $a_{n,\sigma(n)} = 0$. Ainsi $a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} = 0$. Ceci nous donne

$$\begin{array}{rcl} \det(A) & = & \sum_{\sigma \in G} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} + \sum_{\sigma \in S_n \backslash G} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \\ & = & \sum_{\sigma \in G} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n-1),n-1} a_{nn} = a_{nn} \sum_{\sigma \in S_{n-1}} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n-1),n-1} \\ & = & a_{nn} \det(M) = a_{nn} m_{nn}(A). \end{array}$$

La preuve s'achève.

7.3.3. Corollaire. Soit $A = (a_{ij}) \in M_n(K)$. Si la i-ième ligne de A est de la forme $(0, \ldots, 0, a_{ij}, 0, \ldots, 0)$, alors $\det(A) = a_{ij} c_{ij}(A)$.

Démonstration. Soit B la matrice obtenue à partir de A par $n-i$ échanges de lignes $L_i \leftrightarrow L_{i+1}, L_{i+1} \leftrightarrow L_{i+2}, \ldots, L_{n-1} \leftrightarrow L_n$, et $n-j$ échanges de colonnes $C_j \leftrightarrow C_{j+1}, C_{j+1} \leftrightarrow C_{j+2}, \ldots, C_{n-1} \leftrightarrow C_n$. On voit que la n-ième ligne de B est $(0, \ldots, 0, a_{ij})$ et $m_{nn}(B) = m_{ij}(A)$. D'après le lemme 7.2.12, $\det(B) = a_{ij} m_{ij}(B) = a_{ij} m_{ij}(A)$. En outre, d'après la proposition 7.2.5, $\det(B) = (-1)^{n-i+n-j} \det(A)$. Par conséquent,

$$\det(A) = (-1)^{i+j} \det(B) = a_{ij} (-1)^{i+j} m_{ij}(A) = a_{ij} c_{ij}(A).$$

Ceci achève la démonstration.

7.3.4. Expansion de Laplace. Si $A = (a_{ij}) \in M_n(K)$ avec $n \geq 2$, alors

$$\begin{aligned} \det(A) &= a_{i1}c_{i1}(A) + \cdots + a_{in}c_{in}(A), \quad && \textit{développement suivant la } i\textit{-ème ligne.} \\ &= a_{1j}\, c_{1j}(A) + \cdots + a_{nj}\, c_{nj}(A), \quad && \textit{développement suivant la } j\textit{-ème colonne.} \end{aligned}$$

Démonstration. On ne montre que la première égalité. Fixons un indice i avec $1 \leq i \leq n$. Si A_i est la i-ième ligne de A, alors $A_i = A_{i1} + A_{i2} + \cdots + A_{in}$, où $A_{ij} = (0, \ldots, 0, a_{ij}, 0, \ldots, 0)$. Si B_j est la matrice obtenue à partir de A en remplaçant A_i par A_{ij}, alors $\det(B_j) = a_{ij}c_{ij}(B)$ et $c_{ij}(B_j) = c_{ij}(A)$, $j = 1, \ldots, n$. D'après le lemme 7.2.7,

$$\det(A) = \det(B_1) + \cdots + \det(B_n) = a_{i1}c_{i1}(A) + \cdots + a_{in}c_{in}(A).$$

Ceci achève la démonstration.

Le résultat suivant est très utile dans le calcul pratique.

7.3.5. Proposition. Si A et B sont des matrices carrées sur K, alors

$$(1) \quad \det\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det(A)\det(B); \qquad (2) \quad \det\begin{pmatrix} A & 0 \\ C & B \end{pmatrix} = \det(A)\det(B).$$

Démonstration. Écrivons $A = (a_{ij})_{n\times n}$, et posons

$$D = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}.$$

Pour vérifier la première égalité, on dévelope $\det(D)$ suivant la première colonne comme suit:

$$\det(D) = \sum\nolimits_{i=1}^{n} a_{i1}(-1)^{i+1}m_{i1}(D).$$

On procède par récurrence sur n. Si $n = 1$, alors $A = (a_{11})$ et $m_{i1}(D) = \det(B)$. D'où $\det(D) = a_{11}\det(B) = \det(A)\det(B)$. Supposons que $n > 1$. Pour tout $1 \leq i \leq n$, on a

$$m_{i1}(D) = \det\begin{pmatrix} M_{i1} & C_i \\ 0 & B \end{pmatrix},$$

où M_{i1} est la sous-matrice de A obtenue en supprimant la première colonne et la i-ième ligne et C_i est la sous-matrice de C obtenue en supprimant la i-ième ligne. Ainsi $m_{i1}(A) = \det(M_{i1})$, et d'après l'hypothèse de récurrence, $m_{i1}(D) = \det(M_{i1})\det(B) = m_{i1}(A)\det(B)$. Donc

$$\det(D) = \sum\nolimits_{i=1}^{n} a_{i1}(-1)^{i+1}\, m_{i1}(A)\det(B) = \left(\sum\nolimits_{i=1}^{n} a_{i1}c_{i1}(A)\right)\det(B) = \det(A)\det(B).$$

Ceci achève la démonstration.

7.4. Matrices inversibles et règle de Cramer

Dans cette section, on donne des applications de déterminant. Premièrement on étudie l'inversiblité d'une matrice carrée en terme de son déteminant. Ensuite, on donne une méthode pour calculer l'inverse d'une matrice inversible et applique celle-ci pour résoudre un système d'équations linéaires dont la matrice des coefficients est inversible. Enfin, on caractérise le rang d'une matrice en fonction d'ordres de ses mineurs.

7.4.1. Théorème. Si $A \in M_n(K)$, alors A est inversible si et seulement si $\det(A)$ est non nul, et dans ce cas, $\det(A^{-1}) = (\det(A))^{-1}$.

Démonstration. Si A est inversible, alors $AA^{-1} = I_n$. D'après le théorème 7.2.12,

$$\det(A)\det(A^{-1}) = \det(A \cdot A^{-1}) = \det(I_n) = 1.$$

Donc $\det(A) \neq 0$ et $\det(A^{-1}) = (\det(A))^{-1}$.

Réciproquement, si A est non inversible, alors les colonnes $A_1, \cdots, A_n$ de A sont linéairement dépendantes. Ainsi, il existe un $1 \leq j \leq n$ tel que $A_j = \sum_{i \neq j} a_i A_i,\ a_i \in K$. D'après le théorème 7.2.9 et le corollaire 7.2.8, on a

$$\begin{array}{rcl} \det(A) & = & \det(A_1 \cdots A_{j-1}, A_j - \sum_{i \neq j} a_i A_i, A_{j+1}, \cdots, A_n) \\ & = & \det(A_1 \cdots A_{j-1}, 0, A_{j+1}, \cdots, A_n) = 0. \end{array}$$

Ceci achève la démonstration.

7.4.2. Corollaire. Soient $A, B \in M_n(K)$. Si $AB = I$, alors A est inversible et $A^{-1} = B$.

Démonstration. Comme $1 = \det(I_n) = \det(AB) = \det(A)\det(B)$, on a $\det(A) \neq 0$. D'après le théorème 4.1, A est inversible. Or $B = I \cdot B = (A^{-1}A)B = A^{-1}(AB) = A^{-1} \cdot I = A^{-1}$. Ceci achève la démonstration.

7.4.3. Définition. Soit $A \in M_n(K)$ avec $n \geq 2$. La matrice carrée d'ordre n dont le (i, j)-terme est le (i, j)-cofacteur de A s'appelle la *matrice des cofacteurs* de A, notée $\mathrm{cof}(A)$. En outre, la transposée de $\mathrm{cof}(A)$ s'appelle la *matrice adjointe* de A, notée $\mathrm{adj}(A)$.

7.4.4. Proposition. Si $A = (a_{ij}) \in M_n(K)$ avec $n \geq 2$, alors

$$A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) \cdot I_n.$$

Démonstration. Posons $A \cdot \mathrm{adj}(A) = (d_{ij})_{n \times n}$, où

$$d_{ij} = (a_{i1}\ \cdots a_{in}) \begin{pmatrix} c_{j1}(A) \\ \vdots \\ c_{jn}(A) \end{pmatrix} = a_{i1}c_{j1}(A) + \cdots + a_{in}c_{jn}(A).$$

D'après le théorème 7.3.4, on a $d_{ii} = \det(A)$, pour tout $i \leq i \leq n$. Maintenant, on se fixe i, j avec $i \neq j$. Soit B la matrice obtenue à partir de A en remplaçant la j-ième ligne par la i-ième ligne. Comme A, B ont les mêmes lignes sauf la j-ième, $c_{jk}(B) = c_{jk}(A)$ pour tout $1 \leq k \leq n$. Comme B a deux lignes identiques, $\det(B) = 0$. En développant $\det(B)$ suivant la j-ième ligne de B qui est $(a_{i1}, \cdots a_{in})$, on obtient

$$0 = \det(B) = a_{i1}c_{j1}(B) + \cdots + a_{in}c_{jn}(B) = a_{i1}c_{j1}(A) + \cdots + a_{in}c_{jn}(A).$$

D'où, $d_{ij} = 0$ lorsque $i \neq j$. C'est-à-dire, $A \cdot \mathrm{adj}(A) = \det(A) \cdot I_n$. De même, on peut montrer que $\mathrm{adj}(A) \cdot A = \det(A) \cdot I_n$. Ceci achève la démonstration.

7.4.5. Théorème. Soit $A \in M_n(K)$ avec $n \geq 2$. Si A est inversible, alors

$$A^{-1} = (\det(A)^{-1}\,\mathrm{adj}(A).$$

Démonstration. D'après la proposition 7.4.4, on a

$$A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \det(A) \cdot I_n.$$

Si A est inversible, alors $\det(A) \neq 0$. Ainsi

$$A\left((\det(A))^{-1}\operatorname{adj}(A)\right) = (\det(A))^{-1}(A\operatorname{adj}(A)) = (\det(A))^{-1}\det(A)\, I = I.$$

D'où, $A^{-1} = (\det(A))^{-1}\operatorname{adj}(A)$. Ceci achève la démonstration.

Remarque. La formule donnée dans le théorème 7.4.5 pour le calcul de l'inverse d'une matrice inversible s'appelle la *méthode des cofacteurs*.

Exemple. Supposons que la matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

est inversible. Trouver l'inverse de A.
Solution. Par hypothèse, $\det(A) = ad - bc \neq 0$. D'après la méthode de cofacteurs,

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

7.4.6. Règle de Cramer. Soit $AX = B$ un système d'équations linéaires avec A inversible. Pour tout $1 \leq j \leq n$, soit A_j la matrice obtenue à partir de A en remplaçant la j-ième colonne par B. Alors la seule solution de ce système est donnée par

$$x_j = \frac{\det(A_j)}{\det(A)}, \qquad j = 1, 2, \ldots, n.$$

Démonstration. Pour tout $1 \leq j \leq n$, comme A et A_j ont les mêmes colonnes sauf la j-ième, on a $c_{ij}(A) = c_{ij}(A_j)$, $i = 1, 2, \ldots, n$. Posons $B = (b_i)_{n\times 1}$. Alors la seule solution du système est

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A^{-1}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{\det(A)}\begin{pmatrix} c_{11}(A) & c_{21}(A) & \cdots & c_{n1}(A) \\ c_{12}(A) & c_{22}(A) & \cdots & c_{n2}(A) \\ \vdots & \vdots & \ddots & \vdots \\ c_{1n}(A) & c_{2n}(A) & \cdots & c_{nn}(A) \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Par conséquent, pour tout $1 \leq j \leq n$,

$$x_j = \frac{b_1c_{1j}(A) + \cdots + b_nc_{nj}(A)}{\det(A)} = \frac{b_1c_{1j}(A_j) + \cdots + b_nc_{nj}(A_j)}{\det(A)} = \frac{\det(A_j)}{\det(A)}.$$

Ceci achève la démonstration.

Pour conclure cette section, on étudiera le rang d'une matrice en fonction de ses mineurs non nuls.

7.4.7. Lemme. Soit $A = (a_{ij})_{m\times n}$ une matrice sur K. Si les colonnes de A sont linéairement indépendantes, alors A a une sous-matrice inversible d'ordre n.

Démonstration. Soient $\{C_1, \cdots, C_n\}$ les colonnes et $\{L_1, \ldots, L_m\}$ les lignes de A. Si $\{C_1, \cdots, C_n\}$ est libre, alors elle est une base de $\mathcal{C}(A)$. En particulier, dim$\mathcal{C}(A) = n$, et donc dim$\mathcal{L}(A) = n$. Ainsi $\mathcal{L}(A)$ a une base $\{L_{i_1}, L_{i_2}, \cdots, L_{i_n}\}$ avec $1 \le i_1 < i_2, \cdots < i_n \le m$. Or

$$B = \begin{pmatrix} a_{i_1,1} & a_{i_1,2} & \cdots & a_{i_1,n} \\ a_{i_2,1} & a_{i_2,2} & \cdots & a_{i_2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_n,1} & a_{i_n,2} & \cdots & a_{i_n,n} \end{pmatrix}$$

est une sous-matrice carrée de A d'ordre n, qui est inversible puisque les lignes sont linéairement indépendantes. Ceci achève la démonstration.

Le résultat suivant dit que le rang d'une matrice est égal á l'ordre maximal de ses mineurs non nuls.

7.4.8. Théorème. Soient A une matrice non nulle sur K et r un entier positif. Alors rg$(A) = r$ si, et seulement si, A admet un mineur non nul d'ordre r et tous ses mineurs d'ordre $> r$ sont nuls.

Démonstration. Considérons $\mathcal{C}(A) =< A_1, \ldots, A_n >$ de A, où A_j est la j-ième colonne de A, $j = 1, \ldots, n$. Supposons que rg$(A) = r$. Alors $\{A_1, \ldots, A_n\}$ contient une base $\{A_{j_1}, \ldots, A_{j_r}\}$ de $\mathcal{C}(A)$. Or $A' = (A_{j_1} \ \cdots A_{j_r})$ est une sous-matrice de A dont les colonnes sont linéiarement indépendantes. D'après le lemme 7.4.7, A' a une sous-matrice inversible B d'ordre r. Remarquons B est aussi une sous-matrice de A. Ainsi det(B) est un mineur non nul de A d'ordre r.

En outre, soit $C = (C_1, \ \cdots \ , \ C_s)$ une sous-matrice carrée d'ordre s de A avec $s > r$. Alors il existe des indices $k_1, \ldots, k_s$ avec $1 \le k_1 < \cdots < k_s \le n$ tels que C est obtenue à partir de la matrice $(A_{k_1} \ \cdots \ A_{k_s})$ en supprimant $m-s$ lignes. Comme $s > r$, on voit que $\{A_{k_1}, \cdots, A_{k_s}\}$ est liée. Il en de même pour $\{C_1, \ldots, C_s\}$. Ainsi C est non inversible, et donc det$(C) = 0$. Ceci montre la nécessité.

Réciproquement supposons qu'il existe un mineur non nul d'ordre r de A et tous les mineurs d'ordre $> r$ de A sont nuls. Supposons que rg$(A) = t$. D'après la nécessité, A admet un mineur non nul d'ordre t, et donc $t \le r$. En outre, comme tout mineur de A d'ordre $> t$ est nul, on a $r \le t$. Donc $t = r$. Ceci achève la démonstration.

Remarque. Si $A = \text{diag}\{a_1, \ldots, a_n\}$, alors rg$(A)$ est égal au nombre des termes diagonaux non nuls.

7.5. Exercices

1. Soient $n \ge 2$ et $\{\sigma_1, \ldots, \sigma_m\}$ les n-permutations paires. Montrer que si (ab) est une transposition, alors $\{(ab)\sigma_1, \ldots, (ab)\sigma_m\}$ sont les n-permutations impaires. En déduire que $m = \frac{n!}{2}$.

2. Décomposer la permutation suivante en produit de transpositions:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 7 & 5 & 4 & 2 & 6 & 1 & 9 & 3 & 8 \end{pmatrix}.$$

3. Soit $A = (a_{ij})_{6\times 6}$ une matrice carrée d'ordre 6. Dans chacun des cas suivants, déterminer si le terme apparaît dans le déterminant de A ou non; et si oui, trouver son signe.

   (1) $a_{36}a_{54}a_{13}a_{42}a_{65}a_{21}$;

   (2) $a_{62}a_{26}a_{43}a_{11}a_{54}a_{35}$;

   (3) $a_{35}a_{13}a_{51}a_{24}a_{46}a_{63}$.

4. Pour toute $A = (a_{ij})_{n\times n} \in M_n(\mathbb{C})$, on définit $\bar{A} = (\bar{a}_{ij})_{n\times n}$, appelé le *conjugué* de A. On dit que A est *hermitienne* si $A = \overline{A^T}$.

   (1) Montrer que $\overline{\det(A)} = \det(\bar{A})$.

   (2) Si A est hermitienne, montrer que $\det(A)$ est réel.

5. Soit ω une racine cubique de l'unité. Calculer les déterminants sur $\mathbb{C}$ suivants:

$$(1)\ \begin{vmatrix} 1 & 1 & \omega \\ 1 & 1 & \omega^2 \\ \omega^2 & \omega & 1 \end{vmatrix}, \qquad (2)\ \begin{vmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{vmatrix}.$$

6. Calculer les déterminants sur $\mathbb{Q}$ suivants:

$$(1)\ \begin{vmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 3 \end{vmatrix}, \qquad (2)\ \begin{vmatrix} 1 & 1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 1 & 3 \\ 1 & 1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 1 & 0 \\ 2 & 1 & -1 & 1 & 0 \end{vmatrix}.$$

7. Calculer les déterminants sur $\mathbb{Z}_5$ suivants:

$$(1)\ \begin{vmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 1 & 2 \end{vmatrix}, \qquad (2)\ \begin{vmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 1 & 3 \\ 1 & 2 & 3 & 1 & 4 \\ 0 & 2 & 1 & 1 & 2 \\ 3 & 1 & 0 & 1 & 0 \end{vmatrix}.$$

8. Pour quelles valeurs rationnelles de λ, le déterminant suivant est nul?

$$\begin{vmatrix} \lambda-2 & 1 & 1 \\ 1 & \lambda-2 & 1 \\ 1 & 1 & \lambda-2 \end{vmatrix}.$$

9. Montrer que

$$\begin{vmatrix} ta_1+b_1 & tb_1+c_1 & tc_1+a_1 \\ ta_2+b_2 & tb_2+c_2 & tc_2+a_2 \\ ta_3+b_3 & tb_3+c_3 & tc_3+a_3 \end{vmatrix} = (t^3+1)\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

10. Pour quelle valeur de n, la fonction du déterminant

$$\det : M_n(K) \mapsto K : A \mapsto \det(A),$$

est-elle une forme linéaire sur le K-espace vectoriel $M_n(K)$?

11. Calculer le déterminant suivant:

$$\Delta_n = \begin{vmatrix} a_n & a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_1 & a_0 \\ -1 & x & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & x & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & x & 0 \\ 0 & 0 & 0 & 0 & \cdots & -1 & x \end{vmatrix}.$$

12. Calculer le déterminant de chacune des matrices carrées d'ordre n suivantes:

$$(1) \quad \begin{pmatrix} 2 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 2 \end{pmatrix}, \qquad (2) \quad \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 0 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}.$$

13. Soit $n > 2$. Calculer le déterminant suivant:

$$\Delta_n = \begin{vmatrix} \alpha_1 - \beta_1 & \alpha_1 - \beta_2 & \cdots & \alpha_1 - \beta_n \\ \alpha_2 - \beta_1 & \alpha_2 - \beta_2 & \cdots & \alpha_2 - \beta_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n - \beta_1 & \alpha_n - \beta_2 & \cdots & \alpha_n - \beta_n \end{vmatrix}.$$

(*Indication*: Effectuer premièrement l'opération $C_1 - C_2$ et ensuite l'opération $C_2 - C_3$.)

14. Soit $n \geq 2$ un entier. Montrer, à l'aide du déterminant de Vandermonde, que

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n \\ 1 & 2^2 & \cdots & n^2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2^{n-1} & \cdots & n^{n-1} \end{vmatrix} = \Pi_{1 \leq k < n}\, k!$$

15. Montrer que l'application

$$T : \mathbb{R}_n[x] \to \mathbb{R}^n : f \mapsto (f(0), f(1), \ldots, f(n-1))$$

est un isomorphisme d'espaces vectoriels réels. *Indication*: Utiliser le déterminant de Vandemonde pour vérifier la surjectivité de T.

16. Calculer le déterminant de la matrice réelle suivante:

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

Pour quelles valeurs de $a, b, c, d \in \mathbb{R}$, la matrice A est-elle inversible?

17. Considérer la matrice suivante:
$$A = \begin{pmatrix} 0 & a & 0 & b \\ a & 0 & b & 0 \\ 0 & c & 0 & d \\ c & 0 & d & 0 \end{pmatrix}.$$
Pour quelles valeurs de a, b, c et d, la matrice A est-elle inversible?

18. Pour quelles valeurs de a, la famille
$$\{(0, a, 0, 0),\ (1, 0, -1, 0),\ (0, -1, 0, 1 + a),\ (1, 0, 2, -1)\},$$
est-elle une base de $\mathbb{R}^4$?

19. Soient $a, b \in K$. Calculer le déterminant de la matrice carrée d'ordre n suivante:
$$A = \begin{pmatrix} a & b & \cdots & b & b \\ b & a & \cdots & b & b \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & \cdots & b & a \end{pmatrix}.$$
Pour quelles valeurs de a et b, la matrice A, est-elle inversible?

20. Considérer une matrice de Vandermonde suivante:
$$V_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix}.$$
Trouver la condition pour que V_n soit inversible.

21. Montrer, pour tout $\alpha \in \mathbb{R}$, que la matrice
$$A = \begin{pmatrix} \alpha & 1 & -1 & 1 \\ -1 & \alpha & -1 & -1 \\ 1 & 1 & \alpha & -1 \\ -1 & 1 & 1 & \alpha \end{pmatrix}$$
est inversible et trouver A^{-1}. *Indication*: Calculer AA^T.

22. Montrer qu'une matrice complexe antisymétrique d'ordre impaire est non inversible.

23. Soient A et B deux matrices carrées de même type. Montrer que si AB est inversible, alors A et B sont toutes inversibles.

24. Pour quelles valeurs réelles de a, la matrice suivante est inversible?
$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & a \end{pmatrix}.$$

25. Si A, B sont des matrices carrées sur K, montrer qu'une matrice partagée

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

est inversible si et seulement si A et B sont toutes inversibles, et dans ce cas,

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}CB^{-1} \\ 0 & B^{-1} \end{pmatrix}.$$

26. Trouver l'inverse de la matrice suivante par la méthode des cofacteurs.

$$\begin{pmatrix} 3+i & -1 & 0 \\ -6 & 0 & 1+i \\ 0 & 4+i & 5 \end{pmatrix}.$$

27. Soit $A \in M_n(K)$ avec $n \geq 2$. Montrer que $\det(\operatorname{adj}(A)) = (\det(A))^{n-1}$. En déduire que A est inversible si, et seulement si, $\operatorname{adj}(A)$ est inversible.

28. Soient $A, B \in M_2(K)$. Si $\det(B) = 1$, montrer que $\operatorname{tr}(A)\operatorname{tr}(B) = \operatorname{tr}(AB) + \operatorname{tr}(AB^{-1})$.

29. Soit A une matrice carrée sur $\mathbb{Z}$. Montrer que A est inversible sur $\mathbb{Z}$ si, et seulement si, $\det(A) = 1$.

30. Résoudre le système suivant par la règle de Cramer.

(1)
$$\begin{array}{rcrcl} x\cos\theta & - & y\sin\theta & = & a \\ x\sin\theta & + & y\cos\theta & = & b. \end{array}$$

(2)
$$\begin{array}{rcrcrcl} 3x & - & y & + & 2z & = & a \\ -x & + & 2y & - & 3z & = & b \\ 2x & + & 2y & + & z & = & c. \end{array}$$

31. Discuter pour quelles valeurs de $\alpha \in \mathbb{R}$, le système suivant admet une seule solution, et dans ce cas trouver celle-ci en utilisant la règle de Cramer.

$$\begin{array}{rcrcrcl} x & + & y & + & (1-2\alpha)z & = & 2(1+\alpha) \\ (1+\alpha)x & - & (1+\alpha)y & + & (2+\alpha)z & = & 0 \\ 2x & - & 2\alpha y & + & 3z & = & 2(1+\alpha). \end{array}$$

32. Considérer la matrice réelle comme suit:

$$A = \begin{pmatrix} a & 2 & -1 & b \\ 3 & 0 & 1 & -4 \\ 5 & 4 & -1 & 2 \end{pmatrix}.$$

Montrer que $\operatorname{rg}(A) \geq 2$. Pour quelles valeurs de a et b, a-t-on $\operatorname{rg}(A) = 2$?

33. Si $A \in M_n(K)$ est inversible, montrer qu'il existe une sous-matrice inversible d'ordre $n-1$ de A.

34. Soit $A \in M_{m\times n}(K)$. Montrer que $\operatorname{rg}(A) = r$ si, et seulement si, il existe un mineur non nul d'ordre r de A et tout mineur d'ordre $r+1$ de A est nul.

35. Si B est une sous-matrice d'une matrice A sur K, montrer que $\operatorname{rg}(B) \leq \operatorname{rg}(A)$.

36. Montrer que la matrice suivante est de rang plus grand ou égal à $n-1$:

$$\begin{pmatrix} \lambda_1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Pour quelles valeurs de $\lambda_1, \cdots, \lambda_n$ la matrice, est-elle inversible?

Chapitre VIII: Réductions des endomorphismes

Partout dans ce chapitre, on se fixe K un corps et E un K-espace vectoriel de dimension finie. L'objet de ce chapitre est de trouver une base de E dans laquelle la matrice d'un endomorphisme de E est la plus simple possible. D'après le théorème 5.4.4 et la proposition 5.4.7, le problème est équivalent à celui-ci de trouver une matrice la plus simple semblable à une matrice carrée donnée. On discutera premièrement la diagonalisation et ensuite la forme de Jordan de matrices carrées.

8.1. Valeurs propres et vecteurs propres

D'abord, supposons qu'il existe une base $\{u_1, \ldots, u_n\}$ de E dans laquelle la matrice de T est $\text{diag}\{\lambda_1, \ldots, \lambda_n\}$, c'est-à-dire, $T(u_i) = \lambda_i u_i, i = 1, \ldots, n$. Cette observation conduit à la définition suivante.

8.1.1. Définition. (1) Soit $T \in \text{End}(E)$. On dit que $\lambda_0 \in K$ est *valeur propre* de T s'il existe un vecteur non nul $u_0 \in E$ tel que $T(u_0) = \lambda_0 u_0$. Dans ce cas, u_0 s'appelle un *vecteur propre* de T associé à λ_0.

(2) Soit $A \in M_n(K)$. On dit que $\lambda_0 \in K$ est *valeur propre* de A s'il existe un vecteur non nul $u_0 \in K^{(n)}$ tel que $Au_0 = \lambda_0 u_0$. Dans ce cas, u_0 s'appelle un *vecteur propre* de A associé à λ_0.

Exemple. (1) Comme $\mathbb{1}_E(u_0) = u_0$ pour tout $u_0 \in E$, on voit que 1 est une valeur propre de $\mathbb{1}_E$ et tout vecteur non nul de E est un vecteur propre de $\mathbb{1}_E$ associé à 1.

(2) Considérons la rotation ρ_θ du plan d'angle θ avec $\theta \neq k\pi$ pour tout $k \in \mathbb{Z}$. Alors aucun vecteur non nul du plan n'est transformé par ρ_θ en un multiple de lui-même. Donc ρ_θ n'a pas de vecteurs propres dans ce cas.

8.1.2. Proposition. Soit $\{u_1, \ldots, u_n\}$ une base de E. Soient $T \in \text{End}(E)$ et $A = [T]_{\{u_1,\ldots,u_n\}}$. Alors $\lambda_0 \in K$ est une valeur propre de T si et seulement si λ_0 est une valeur de A, et dans ce cas, $u_0 \in E$ est un vecteur propre de T associé à λ_0 si et seulement si $P^{\{u_0\}}_{\{u_1,\ldots,u_n\}}$ est un vecteur propre de A associé à λ_0.

Démonstration. Soient $\lambda_0 \in K$ et $u_0 \in E$. Alors λ_0 est une valeur propre de T et u_0 est un vecteur propre asscoicé à λ_0 si et seulement si $u_0 \neq 0$ et $T(u_0) = \lambda_0 u_0$ si et seulement si $P^{\{u_0\}}_{\{u_1,\ldots,u_n\}} \neq 0$ et $P^{\{T(u_0)\}}_{\{u_1,\ldots,u_n\}} = P^{\{\lambda_0 u_0)\}}_{\{u_1,\ldots,u_n\}}$ si et seulement si $P^{\{u_0\}}_{\{u_1,\ldots,u_n\}} \neq 0$ et $AP^{\{u_0\}}_{\{u_1,\ldots,u_n\}} = \lambda_0 P^{\{u_0\}}_{\{u_1,\ldots,u_n\}}$ si et seulement si λ_0 est une valeur propre de A et $P^{\{u_0\}}_{\{u_1,\ldots,u_n\}}$ est un vecteur propre de A associé à λ_0. Ceci achève la démonstration.

8.1.3. Définition. (1) On dit que $T \in \text{End}(E)$ est *diagonalisable* s'il existe une base de E dans laquelle la matrice de T est diagonale.

(2) On dit que $A \in M_n(K)$ est *diagonalisable sur* K s'il existe une matrice inversible $P \in M_n(K)$ telle que $P^{-1}AP$ est diagonale.

8.1.4. Théorème. Soit $\{u_1, \ldots, u_n\}$ une base de E. Soient $T \in \text{End}(E)$ et $A = [T]_{\{u_1,\ldots,u_n\}}$. Les conditions suivantes sont équivalentes:

(1) T est diagonalisable.

(2) T admet n vecteurs propres linéairement indépendants.

(3) A est diagonalisable sur K.

(4) A admet n vecteurs propres linéairement indépendants dans $K^{(n)}$.

Démonstration. Supposons que T est diagonalisable. Alors il existe une base $\{v_1, \ldots, v_n\}$ de E telle que $[T]_{\{v_1,\ldots,v_n\}} = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$, c'est-à-dire, $T(v_i) = \lambda_i v_i$, $i = 1, \ldots, n$. Ainsi $v_1, \ldots, v_n$ sont n vecteurs propres linéairement indépendants de T. Ceci montre que (1) implique (2).

Ensuite, supposons que $v_1, \ldots, v_n$ sont n vecteurs propres linéairement indépendants de T. Alors $T(v_i) = \lambda_i v_i$, $\lambda_i \in K$, $i = 1, \cdots, n$. Comme $\dim(E) = n$, on voit que $\{v_1, \ldots, v_n\}$ est une base de E telle que $[T]_{\{v_1,\ldots,v_n\}} = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$. Soit P la matrice de passage de $\{u_1, \ldots, u_n\}$ vers $\{v_1, \ldots, v_n\}$. Alors $P^{-1}AP = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$, c'est-à-dire, A est diagonalisable sur K. Ceci montre que (2) implique (3).

Supposons que A est diagonalisable sur K. Alors il existe $P \in M_n(K)$ inversible telle que $P^{-1}AP = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$. C'est-à-dire, $AP = P\text{diag}\{\lambda_1, \ldots, \lambda_n\}$. Partageons $P = (P_1, \ldots, P_n)$ en colonnes. Alors $(AP_1, \cdots, AP_n) = (\lambda_1 P_1, \cdots, \lambda_n P_n)$. D'où, $AP_i = \lambda_i P_i$, $i = 1, \ldots, n$. Donc $P_1, \cdots, P_n$ sont n vecteur propres de A, qui sont linéairement indépendants puisque P est inversible. Ceci montre que (3) implique (4).

Enfin, supposons que A a n vecteurs propres linéairement indépendants $P_1, \ldots, P_n$ dans $K^{(n)}$. Alors $AP_i = \lambda_i P_i$, $\lambda_i \in K$, $i = 1, \ldots, n$. Or $P = (P_1, \cdots, P_n)$ est inversible telle que $AP = P\text{diag}\{\lambda_1, \cdots, \lambda_n\}$. C'est-à-dire, $P^{-1}AP = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$. Posons $(v_1, \cdots v_n) = (u_1, \ldots, u_n)P$. Alors $\{v_1, \cdots v_n\}$ est une base de E, et la matrice de passage de $\{u_1, \ldots, u_n\}$ vers $\{v_1, \ldots, v_n\}$ est P. Ainsi

$$[T]_{\{v_1,\ldots,v_n\}} = P[T]_{\{u_1,\ldots,u_n\}}P^{-1} = P^{-1}AP = \text{diag}\{\lambda_1, \ldots, \lambda_n\}.$$

Ceci montre que (4) implique (1). La preuve s'achève.

D'après les résultats précédents, il suffit d'étudier la diagonalisation de matrices carrées. Ceci consiste à trouver les valeurs propres et des vecteurs propres associés. Pour trouver les valeurs propres, on introduira le polynôme caractéristique d'une matrice carrée. Pour ce faire, on désigne par $k[\lambda]$ le K-espace vectoriel des polynômes sur K en indéterminée λ. Remarquons que $K[\lambda]$ est muni d'une multiplication qui satisfont à tous les axiomes d'un corps sauf qu'un polynôme non constant n'a aucun inverse dans $K[\lambda]$. Maintenant, on généralise la notion du déterminant de matrices carrées sur K aux matrices carrées sur $K[\lambda]$.

8.1.5. Définition. Soit $A(\lambda) = (a_{ij}(\lambda))_{n\times n}$, où $a_{ij}(\lambda) \in K[\lambda]$. On définit

$$\det(A(\lambda)) = \sum\nolimits_{\sigma \in S_n} \text{sgn}(\sigma)\, a_{1,\sigma(1)}(\lambda) \cdots a_{n,\sigma(n)}(\lambda) \in K[\lambda].$$

Remarque. Toutes les propriétés du déterminant de matrices carrées sur K énoncées dans le chapitre VII demeurent valides pour le déterminant de matrices carrées sur $K[\lambda]$. Cependant, il est interdit d'effectuer les opérations comme $L_i \pm \frac{1}{f(\lambda)} L_j$ ou $C_i \pm \frac{1}{f(\lambda)} C_j$ avec $f(\lambda)$ non constant, parce que $\frac{1}{f(\lambda)} \notin K[\lambda]$.

8.1.6. Définition. Soit $A \in M_n(K)$. Le polynôme

$$\chi_A(\lambda) = \det(A - \lambda I_n)$$

s'appelle *polynôme caractéristique* de A.

8.1.7. Proposition. Si $A \in M_n(K)$, alors

$$\chi_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \text{tr}(A) \lambda^{n-1} + a_2 \lambda^{n-2} + \cdots + a_{n-1}\lambda + \det(A),\ a_i \in K.$$

Démonstration. Si $A = (a_{ij})_{n\times n}$, alors $A - \lambda I_n = (a_{ij}(\lambda))_{n\times n}$, où $a_{ii}(\lambda) = a_{ii} - \lambda$, pour tout $1 \le i \le n$ et $a_{ij}(\lambda) = a_{ij}$, pour tous $i \ne j$. D'après la définition, on a

$$\chi_A(\lambda) = \sum\nolimits_{\sigma \in S_n} \text{sgn}(\sigma)\, a_{1,\sigma(1)}(\lambda) \cdots a_{n,\sigma(n)}(\lambda).$$

Si $\sigma \in S_n$ est non identité, alors $\sigma(r) \neq r$ pour un certain $1 \leq r \leq n$. Posant $s = \sigma(r)$, alors $\sigma(s) \neq s$. Ainsi $a_{r,\sigma(r)}(\lambda) = a_{r,\sigma(r)}$ et $a_{s,\sigma(s)}(\lambda) = a_{s,\sigma(s)}$. D'où, $a_{1,\sigma(1)}(\lambda) \cdots \alpha_{n,\sigma(n)}(\lambda)$ est de degré $\leq n-2$. Ainsi $(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$ est le seul terme de $\chi_A(\lambda)$ qui contient λ^n et λ^{n-1} dont les coefficients sont $(-1)^n$ et $(-1)^{n-1}\mathrm{tr}(A)$, respectivement. Par conséquent,

$$\chi_A(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}\mathrm{tr}(A)\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_{n-1}\lambda + a_n,\ a_i \in K.$$

Enfin, $a_n = \chi_A(0) = \det(A - 0 \cdot I_n) = \det(A)$. Ceci achève la preuve.

8.1.8. Lemme. Si $A, B \in M_n(K)$ sont semblables, alors $\chi_A(\lambda) = \chi_B(\lambda)$.
Démonstration. Soit P inversible telle que $B = P^{-1}AP$. Or

$$\begin{aligned} \det(B - \lambda I_n) &= \det(P^{-1}(A - \lambda I_n)P) \\ &= \det(P^{-1})\det(A - \lambda I_n)\det(P) \\ &= \det(P^{-1})\det(P)\det(A - \lambda I_n) \\ &= \det(A - \lambda I_n)\,. \end{aligned}$$

Ceci achève la démonstration.

8.1.9. Proposition. Soient $A \in M_n(K)$ et $\lambda_0 \in K$. Les conditions suivantes sont équivalentes:
(1) λ_0 est une valeur propre de A.
(2) $\mathrm{rg}\,(A - \lambda_0 I_n) < n$.
(3) λ_0 est racine de $\chi_A(\lambda)$.
Démonstration. La valeur λ_0 est une valeur propre de A si, et seulement si, il existe $u_0 \in K^{(n)}$ non nul tel que $Au_0 = \lambda_0 u_0 = (\lambda_0 I_n)u_0$ si, et seulement si, $(A - \lambda_0 I_n)u_0 = 0$ avec u_0 non nul si, et seulement si, le système homogène $(A - \lambda_0 I_n)X = 0$ admet une solution non nulle si, et seulement si, $\mathrm{rg}(A - \lambda_0 I_n) < n$ si, et seulement si, $\det(A - \lambda_0 I_n) = 0$, c'est-à-dire, $\chi_A(\lambda_0) = 0$. Ceci achève la démonstration.

Remarque. (1) Comme $\chi_A(\lambda)$ est de degré n, A admet au plus n valeurs propres.
(2) La matrice $A - \lambda_0 I_n$ s'appelle *matrice caractéristique* de A associée à λ_0.

Exemple. Si $D = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$, alors $\chi_D(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$. D'où, les valeurs propres de D sont $\lambda_1, \ldots, \lambda_n$.

8.1.10. Définition. Soit λ_0 une valeur propre de A. Comme $\lambda - \lambda_0$ divise $\chi_A(\lambda)$, on appelle *multiplicité algébrique* de λ_0, notée $\mathrm{ma}\,(\lambda_0)$, le plus grand exposant r tel que $(\lambda - \lambda_0)^r$ divise $\chi_A(\lambda)$.

Remarque. D'après la définition, $\mathrm{ma}\,(\lambda_0) = r$ si et seulement si $\chi_A(\lambda) = (\lambda - \lambda_0)^r f(\lambda)$ avec $f(\lambda) \in K[\lambda]$ tel que $f(\lambda_0) \neq 0$.

Le résultat suivant est trivial.

8.1.11. Proposition. Soit $A \in M_n(K)$ dont λ_0 est une valeur propre.
(1) Un vecteur $u_0 \in K^{(n)}$ est un vecteur propre de A associé à λ_0 si, et seulement si, u_0 est une solution non nulle du système homogène $(A - \lambda_0 I_n)X = 0$.
(2) $K^{(n)}(A, \lambda_0) = \{u \in K^{(n)} \mid Au = \lambda_0 u\}$ est l'espace-solution du système homgogène $(A - \lambda_0)X = 0$, qui s'appelle l'*espace propre* de A associé à λ_0.

Remarque. D'après la proposition 8.1.11(2), $\dim(K^{(n)}(A, \lambda_0))$ est le nombre maximal de vecteurs propres linéairement indépendants de A associés à λ_0, appelé *multiplicité géométrique* de λ_0 et noté $\mathrm{mg}\,(\lambda_0)$. Rappelons que $\mathrm{mg}(\lambda_0) = n - \mathrm{rg}(A - \lambda_0 I_n)$.

Le résultat suivant nous dit la relation entre la multiplicité algébrique et la multiplicité géométrique de valeur propre.

8.1.12. Théorème. Soit $A \in M_n(K)$. Si λ_0 est une valeur propre de A, alors $\mathrm{mg}\,(\lambda_0) \leq \mathrm{ma}\,(\lambda_0)$.

Démonstration. Soit $T : K^{(n)} \to K^{(n)} : u \mapsto Au$. Alors A est la matrice de T dans la base canonique de $K^{(n)}$. Si $\mathrm{mg}(\lambda_0) = r$, alors $K^{(n)}(A, \lambda_0)$ a une base $\{u_1, \ldots, u_r\}$. Cette dernière se prolonge en une base $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ de $K^{(n)}$. Pour tout $1 \leq j \leq r$, on a $T(u_j) = Au_j = \lambda_0 u_j$. Pour tout $r < j \leq n$, on a

$$T(u_j) = a_{1j}u_1 + \cdots + a_{rj}u_r + a_{r+1,j}u_{r+1} + \cdots + a_{nj}u_n, \quad a_{ij} \in K.$$

D'où,

$$[T]_{\{u_1,\ldots,u_r,u_{r+1},\ldots,u_n\}} = \begin{pmatrix} \lambda_0 & \cdots & 0 & a_{1,r+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & \lambda_0 & a_{r,r+1} & \cdots & a_{rn} \\ 0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_0 I_r & C \\ 0 & B \end{pmatrix}.$$

D'après le théorème 5.4.4 et le lemme 8.1.8, on a

$$\chi_A(\lambda) = \det \begin{pmatrix} \lambda_0 I_r - \lambda I_r & C \\ 0 & B - \lambda I_{n-r} \end{pmatrix} = (\lambda_0 - \lambda)^r \chi_B(\lambda).$$

Par conséquent, $\mathrm{ma}\,(\lambda_0) \geq r = \mathrm{mg}\,(\lambda_0)$. Ceci achève la démonstration.

Remarque. Si $\mathrm{ma}\,(\lambda_0) = 1$, alors $\mathrm{mg}\,(\lambda_0) = \mathrm{ma}\,(\lambda_0) = 1$.

8.1.11. Théorème. Soit $A \in M_n(K)$. Si $\lambda_1, \ldots, \lambda_r$ sont des valeurs propres distinctes de A, alors

$$\sum_{i=1}^{r} K^{(n)}(A, \lambda_i) = \oplus_{i=1}^{r} K^{(n)}(A, \lambda_i).$$

Par conséquent, si $\mathcal{U}_i$ est une famille libre de vecteurs propres de A associés à λ_i, alors $\mathcal{U}_1 \cup \cdots \cup \mathcal{U}_r$ est libre.

Démonstration. Le résultat est trivial si $r = 1$. Supposons que $r > 1$ et le résultat est vrai pour $r - 1$, c'est-à-dire, $\sum_{i=1}^{r-1} K^{(n)}(A, \lambda_i) = \oplus_{i=1}^{r-1} K^{(n)}(A, \lambda_i)$. Soit $u_1 + \cdots + u_{r-1} + u_r = 0,\ u_i \in K^{(n)}(A, \lambda_i)$. Multipliant par λ_r et par A respectivement, on obtient

$$\lambda_r u_1 + \cdots + \lambda_r u_{r-1} + \lambda_r u_r = 0 \quad \text{et} \quad \lambda_1 u_1 + \cdots + \lambda_{r-1} u_{r-1} + \lambda_r u_r = 0.$$

Ainsi $(\lambda_r - \lambda_1)u_1 + \cdots + (\lambda_r - \lambda_{r-1})u_{r-1} = 0$. Par hypothèse de récurrence, $(\lambda_r - \lambda_i)u_i = 0$ pour tout $1 \leq i < r$. Ainsi $u_i = 0$ car $\lambda_r - \lambda_i \neq 0$ pour tout $1 \leq i < r$. Donc $u_r = 0$. Ceci montre que la somme $\sum_{i=1}^{r} K^{(n)}(A, \lambda_i)$ est directe. La preuve s'achève.

8.2. Diagonalisation

Cette section a pour but de discuter quand et comment on peut diagonaliser une matrice carrée et un endomorphisme. On commence par un lemme qui suit du Théorème 7.4.8.

8.2.1. Lemme. Le rang d'une matrice diagonale est égal au nombre de termes diagonaux non nuls.

8.2.2. Lemme. Soit A une matrice semblable à $D = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$.

(1) $\chi_A(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$. Par conséquent, $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de A, en comptant les multiplicités algébriques.

(2) $\mathrm{mg}\,(\lambda_i) = \mathrm{ma}\,(\lambda_i)$, $i = 1, \cdots, n$.

Démonstartion. D'abord, $\chi_A(\lambda) = \chi_D(\lambda) = \det(D - \lambda I_n) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$, par le lemme 8.1.8. En conséquence, $\mathrm{ma}\,(\lambda_i) = \#\{j \mid 1 \le j \le n,\ \lambda_j = \lambda_i\}$, pour tout $1 \le i \le n$. D'autre part, comme A est semblable à D, on a $(A - \lambda_i I_n)$ est semblable à $D - \lambda_i I_n$. En particulier, $\mathrm{rg}\,(A - \lambda_i I_n) = \mathrm{rg}\,(D - \lambda_i I_n)$. Par conséquent,

$$\mathrm{mg}\,(\lambda_i) = n - \mathrm{rg}\,(A - \lambda_i I_n) = n - \mathrm{rg}\,(\mathrm{diag}\{\lambda_1 - \lambda_i, \ldots, \lambda_n - \lambda_i\}),$$

ce qui est égal au nombre de termes diagonaux nuls de $\mathrm{diag}\{\lambda_1 - \lambda_i, \ldots, \lambda_n - \lambda_i\}$, c'est-à-dire, le nombre d'indices j avec $1 \le j \le n$ tel que $\lambda_j = \lambda_i$. Donc $\mathrm{mg}\,(\lambda_i) = \mathrm{ma}\,(\lambda_i)$. Ceci achève la démonstration.

On dit que $f(\lambda) \in K[\lambda]$ est *scindé* sur K si $f(\lambda)$ se factorise en produit de facteurs linéaires sur K. For example, $\lambda^2 + 1$ n'est pas scindé sur $\mathbb{R}$, mais il est scindé sur $\mathbb{C}$. En effet, tout polynôme sur $\mathbb{C}$ est scindé sur $\mathbb{C}$.

8.2.3. Théorème. Une matrice $A \in M_n(K)$ est diagonalisable sur K si et seulement si les deux conditions suivantes sont satisfaites.

(1) Le polynôme caractéristique $\chi_A(\lambda)$ est scindé sur K.

(2) Pour toute valeur propre λ_0 de A, $\mathrm{mg}\,(\lambda_0) = \mathrm{ma}\,(\lambda_0)$.

Démonstration. La nécessité suit du lemme 8.2.2. Pour montrer la suffisance, on suppose que les deux conditions sont satisfaites. Alors $\chi_A(\lambda) = (\lambda_1 - \lambda)^{n_1} \cdots (\lambda_r - \lambda)^{n_r}$ avec $\lambda_1, \cdots, \lambda_r \in K$ deux à deux distincts et $n_i > 0$. Ainsi $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres distinctes de A. De plus, pour tout $1 \le i \le r$, on a $\mathrm{mg}(\lambda_i) = \mathrm{ma}(\lambda_i) = n_i$, et donc on peut trouver une famille libre $\mathcal{B}_i$ de n_i vecteurs propres de A associés à λ_i, $i = 1, \cdots, r$. D'après le théorème 8.1.11, $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est une famille libre de vecteurs propres de A. Comme $|\mathcal{B}| = \sum_{i=1}^r n_i = n$, d'après le théorème 8.1.4(4), A est diagonalisable sur K. Ceci achève la démonstration.

8.2.4. Corollaire. Soit $A \in M_n(K)$. Si $\chi_A(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$ avec $\lambda_1, \ldots, \lambda_n \in K$ deux à deux distincts, alors A est diagonalisable sur K.

Démonstration. Comme $\lambda_1, \ldots, \lambda_n$ sont deux à deux dinstincts, $\mathrm{ma}\,(\lambda_i) = 1$, et donc $\mathrm{mg}(\lambda_i) = \mathrm{ma}(\lambda_i)$ pour tout $1 \le i \le n$. D'après le théorème 8.4.3, A est diagonalisable. Ceci achève la démonstration.

8.2.5. Procédé de la diagonalisation d'une matrice carrée A d'ordre n.

(1) Calculer le polynôme caractéristique $\chi_A(\lambda) = \det(A - \lambda I_n)$.

(2) Si $\chi_A(\lambda)$ n'est pas scindé sur K, alors A n'est pas diagonalisable sur K.

(3) Si oui, écrire $\chi_A(\lambda) = (\lambda_1 - \lambda)^{n_1} \cdots (\lambda_r - \lambda)^{n_r}$ avec $\lambda_1, \ldots, \lambda_r$ deux à deux distincts et $n_i > 0$.

(4) Si $n - \mathrm{rg}(A - \lambda_i I_n) \ne n_i$ pour un certain $1 \le i \le r$, alors A n'est pas diagonalisable sur K.

(5) Si oui, trouver une base $\mathcal{B}_i$ de $K^{(n)}(A, \lambda_i)$ en résolvant le système $(A - \lambda_i I_n)X = 0$ pour tout $i = 1, \cdots, r$.

(6) La matrice P formée par les vecteurs de $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est inversible telle que

$$P^{-1}AP = \mathrm{diag}\{\overbrace{\lambda_1, \ldots, \lambda_1}^{n_1}, \ldots, \overbrace{\lambda_r, \ldots, \lambda_r}^{n_r}\}.$$

8.2.6. Procédé de la diagonalisation d'un endomorphisme T de E.
(1) Choisir une base $\{u_1, \ldots, u_n\}$ quelconque de E et calculer $A = [T]_{\{u_1,\ldots,u_n\}}$.
(2) Si A est non diagonalisable sur K, alors T est non diagonalisable.
(3) Si oui, trouver P inversible telle que $P^{-1}AP$ est diagonale.
(4) Si $(v_1, \cdots, v_n) = (u_1, \ldots, u_n)P$, alors $\{v_1, \ldots, v_n\}$ est une base de E dans laquelle la matrice de T est $P^{-1}AP$.

8.3. Polynômes annulateurs

Soit $T \in \text{End}(E)$. Définissons $T^0 = \mathbf{1}_E$ et $T^r = \overbrace{T \cdots T}^{r \text{ fois}} \in \text{End}(E)$ pour tout $r > 0$. En général, on a le définition suivante.

8.3.1. Définition. Soit $f(\lambda) = a_0 + a_1\lambda + \cdots + a_r\lambda^r \in K[\lambda]$.
(1) Pour $A \in M_n(K)$, on pose $f(A) = a_0 I_n + a_1 A + \cdots + a_r A^r \in M_n(K)$.
(2) Pour $T \in \text{End}(E)$, on pose $f(T) = a_0\mathbf{1} + a_1 T + \cdots + a_r T^r \in \text{End}(E)$.

On voit facilement que le résultat suivant est vrai.

8.3.2. Lemme. Soient $A \in M_n(K)$ et $T \in \text{End}(E)$. Pour tous $f(\lambda), g(\lambda), h(\lambda) \in K[\lambda]$, les énoncés suivants sont valides.
(1) Si $f = g + h$, alors $f(A) = g(A) + h(A)$ et $f(T) = g(T) + h(T)$.
(2) Si $f = gh$, alors $f(A) = g(A)h(A)$ et $f(T) = g(T)h(T)$.
(3) $f(A)g(A) = g(A)f(A)$ et $f(T)g(T) = g(T)f(T)$.

Remarque. Considérons $T - \lambda_0\mathbf{1}$ avec $\lambda_0 \in K$. Alors λ_0 est une valeur propre de T si et seulement si $\text{Ker}(T - \lambda_0\mathbf{1})$ est non nul. Dans ce cas, les vecteurs propres de T associés à λ_0 sont les vecteurs non nuls de $\text{Ker}(T - \lambda_0\mathbf{1})$.

On va premièrement appliquer la diagonalisation de matrices carrées au calcul de polynômes matriciels.

8.3.3. Proposition. Si $f(\lambda) \in K[\lambda]$, alors $f(\text{diag}\{\lambda_1, \ldots, \lambda_n\}) = \text{diag}\{f(\lambda_1), \ldots, f(\lambda_n)\}$.
Démonstration. Pour tous $a, \lambda_1, \ldots, \lambda_n; \mu_1, \ldots, \mu_n \in K$, on a

$$a\,\text{diag}\{\lambda_1, \ldots, \lambda_n\} = \text{diag}\{a\lambda_1, \ldots, a\lambda_n\},$$
$$\text{diag}\{\lambda_1, \ldots, \lambda_n\} + \text{diag}\{\mu_1, \ldots, \mu_n\} = \text{diag}\{\lambda_1 + \mu_1, \ldots, \lambda_n + \mu_n\} \text{ et}$$
$$\text{diag}\{\lambda_1, \ldots, \lambda_n\} \cdot \text{diag}\{\mu_1, \ldots, \mu_n\} = \text{diag}\{\lambda_1 \cdot \mu_1, \ldots, \lambda_n \cdot \mu_n\}.$$

Si $f(\lambda) = \sum_{i=0}^r a_i\lambda^i$, alors

$$\begin{aligned} & f(\text{diag}\{\lambda_1, \ldots, \lambda_n\}) = \textstyle\sum_{i=0}^r a_i\,\text{diag}\{\lambda_1, \ldots, \lambda_n\}^i \\ = \;& \textstyle\sum_{i=0}^r a_i\,\text{diag}\{\lambda_1^i, \ldots, \lambda_n^i\} = \sum_{i=0}^r \text{diag}\{a_i\,\lambda_1^i, \ldots, a_i\,\lambda_n^i\} \\ = \;& \text{diag}\{\textstyle\sum_{i=0}^r a_i\,\lambda_1^i,\, \ldots,\, \sum_{i=0}^r a_i\,\lambda_n^i\} = \text{diag}\{f(\lambda_1), \ldots, f(\lambda_n)\}. \end{aligned}$$

Ceci achève la démonstration.

8.3.4. Proposition. Soient $A, B \in M_n(K)$. Si $A = P^{-1}BP$, alors $f(A) = P^{-1}f(B)\,P$, pour tout $f(\lambda) \in K[\lambda]$.
Démonstration. On voit que $A^2 = (P^{-1}BP)(P^{-1}BP) = P^{-1}B^2P$ et

$$A^3 = (P^{-1}B^2P)(P^{-1}BP) = P^{-1}B^3P.$$

En général, $A^i = P^{-1}B^iP$, pour tout $i \geq 0$. Si $f(\lambda) = \sum_{i=0}^r a_i\lambda^i$, alors

$$f(A) = \sum\nolimits_{i=0}^{r} a_iA^i = \sum\nolimits_{i=0}^{r} a_i(P^{-1}B^iP) = P^{-1}(\sum\nolimits_{i=0}^{r} a_iB^i)\,P = P^{-1}f(B)\,P.$$

Ceci achève la démonstration.

Remarque. Si $A = PBP^{-1}$, alors $f(A) = Pf(B)\,P^{-1}$.

8.3.5. Définition. Soient $T \in \mathrm{End}(E)$ et $A \in M_n(K)$. Un polynôme non nul $f(\lambda)$ sur K s'appelle un *annulateur* de T (respectivement, de A) si $f(T) = \mathbf{0}$ (respectivement, $f(A) = 0$).

Remarque. Tout annulateur est de degré ≥ 1.

Exemple. (1) $\lambda - 1$ est un annulateur de $\mathbf{1}_E$ et λ est un annulateur de $\mathbf{0}$.
(2) Considérons un *bloc de Jordan d'ordre n* comme suit:

$$J_n(\lambda_0) = \begin{pmatrix} \lambda_0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & \lambda_0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \lambda_0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \lambda_0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & \lambda_0 \end{pmatrix}_{n\times n}, \quad \lambda_0 \in K.$$

On sait que $(A - \lambda_0 I)^n = 0$. Ainsi $(\lambda - \lambda_0)^n$ est un annulateur de $J_n(\lambda_0)$.

8.3.6. Proposition. Soit $\{u_1, \ldots, u_n\}$ une base de E et soit $T \in \mathrm{End}(E)$. Si $[T]_{\{u_1,\ldots,u_n\}} = A$, alors $[f(T)]_{\{u_1,\ldots,u_n\}} = f(A)$ pour tout $f(\lambda) \in K[\lambda]$.

Démonstration. Posons $f(\lambda) = a_0 + a_1\lambda + \cdots + a_r\lambda^r$. D'après la proposition 5.3.12, $[T^i] = [T]^i = A^i$ pour tout $i \geq 0$. D'après la proposition 5.3.9,

$$\begin{aligned} [f(T)] &= [a_0\mathbf{1} + a_1T + \cdots + a_rT^r] \\ &= \alpha_0[\mathbf{1}] + a_1[T] + \cdots + a_r[T^r] \\ &= a_0I_n + a_1A + \cdots + a_rA^r = f(A). \end{aligned}$$

8.3.7. Lemme. Soient $T \in \mathrm{End}(E)$ et A la matrice de T dans une base de E. Soient $f(\lambda)$ et $g(\lambda)$ des polynômes non nuls sur K.
(1) Si $f(\lambda)$ est un annulateur de T, alors $f(\lambda)g(\lambda)$ l'est.
(2) $f(\lambda)$ est un annulateur de T si et seulement si $f(\lambda)$ est un annulateur de A.

Démonstration. (1) Posons $h(\lambda) = f(\lambda)g(\lambda)$. Si $f(T) = 0$, alors $h(T) = f(T)g(T) = 0$.
(2) $f(T) = 0$ si et seulement si $[f(T)] = 0$ si et seulement si $f([T]) = f(A) = 0$. Ceci achève la démonstration.

Soient $T \in \mathrm{End}(E)$ et A la matrice de T dans une base de E. On définit le *polynôme caractéristique* $\chi_T(\lambda)$ de T comme étant $\chi_A(\lambda)$. Ceci est correctement défini parce que les matrices de T dans deux bases sont semblables, et donc ont le même polynôme caractéristique. On montrera que le polynôme caractéristique est toujours un annulateur. Pour ce faire, on remarque que si $A(\lambda)$ est une matrice carrée d'ordre n de polynômes sur K de degré $\leq r$, alors

$$A(\lambda) = A_r \cdot \lambda^r + \cdots + A_1 \cdot \lambda + A_0, \text{ où } A_0, A_1, \ldots, A_r \in M_n(K).$$

8.3.8. Théorème de Hamilton-Cayley. (1) Si $A \in M_n(K)$, alors $\chi_A(\lambda)$ est un annulateur de A.
(2) Si $T \in \text{End}(E)$, alors $\chi_T(\lambda)$ est un annulateur de T.

Démonstration. Posons $\chi_A(\lambda) = \sum_{i=0}^{n} a_i\lambda^i$, où $a_i \in K$. Soit $B(\lambda) = (c_{ij}(\lambda))_{n\times n}^T$, la matrice adjointe de $(A-\lambda I)$, où $c_{ij}(\lambda)$ est le (i,j)-cofacteur de $A-\lambda I$. Comme $c_{ij}(\lambda)$ est de degré $\leq n-1$, on a $B(\lambda) = \sum_{i=0}^{n-1} B_i\lambda^i$, où $B_i \in M_n(K)$. Remarquons que $(A - \lambda I)B(\lambda) = \det(A - \lambda I) \cdot I = \chi_A(\lambda) \cdot I = \sum_{i=0}^{n} a_i I\lambda^i$. D'autre part,

$$\begin{aligned} (A-\lambda I)B(\lambda) &= (A-\lambda I)\textstyle\sum_{i=0}^{n-1} B_i\lambda^i = \sum_{i=0}^{n-1} AB_i\lambda^i - \sum_{i=0}^{n-1} B_i\lambda^{i+1} \\ &= \textstyle\sum_{i=0}^{n-1} AB_i\lambda^i - \sum_{i=1}^{n} B_{i-1}\lambda^i = AB_0 + \sum_{i=1}^{n-1}(AB_i - B_{i-1})\lambda^i - B_{n-1}\lambda^n. \end{aligned}$$

D'où, $a_0I = AB_0$, $a_iI = AB_i - B_{i-1}$, $i = 1, \ldots, n-1$, et $a_nI = -B_{n-1}$. Ceci nous donne $a_0I = AB_0$, $a_iA^i = A^{i+1}B_i - A^iB_{i-1}$, $i = 1, \ldots, n-1$, et $a_nA^n = -A^nB_{n-1}$. Par conséquent,

$$\begin{aligned} \chi_A(A) &= \textstyle\sum_{i=0}^{n} a_iA^i = AB_0 + \sum_{i=1}^{n-1} A^{i+1}B_i - \sum_{i=1}^{n-1} A_iB_{i-1} - A^nB_{n-1} \\ &= \textstyle\sum_{i=0}^{n-1} A^{i+1}B_i - \sum_{i=1}^{n} A^iB_{i-1} = \sum_{i=0}^{n-1} A^{i+1}B_i - \sum_{i=0}^{n-1} A^{i+1}B_i = 0. \end{aligned}$$

Ceci montre la partie (1), et la partie (2) suit du lemme 8.3.7(2). La preuve s'achève.

Le résultat facile suivant nous donne une autre méthode de calculer l'inverse d'une matrice inversible en utilisant ses annulateurs.

8.3.9. Proposition. Soient $A \in M_n(K)$ et $f(\lambda) = a_0 + a_1\lambda + \cdots + a_r\lambda^r$ un annulateur de A. Si $a_0 \neq 0$, alors A est inversible et

$$A^{-1} = -\frac{a_1}{a_0}I_n - \frac{a_2}{a_0}A - \cdots - \frac{a_n}{a_0}A^{r-1}.$$

8.4. Polynôme minimal

Cette section a pour but de trouver les polynômes annulateurs du plus petit de degré d'une matrice carrée. Cela nous donnera une autre condition nécessaire et suffisante pour qu'une matrice carrée soit diagonalisable.

Un polynôme $f(\lambda) = a_0 + a_1\lambda + \cdots + a_r\lambda^r$ sur K est dit *normalisé* si $a_r = 1_K$.

8.4.1. Définition. Soient $T \in$End(E) et $A \in M_n(K)$. Un polynôme $m(\lambda)$ sur K s'appelle un *polynôme minimal* de T (respectivement, de A) si
(1) $m(\lambda)$ est normalisé,
(2) $m(\lambda)$ est un annulateur de T (respectivement, de A), et
(3) le degré de $m(\lambda)$ est le plus petit parmi les degrés des annulateurs de T (respectivement, de A).

Exemple. (1) $\lambda - 1$ est un polynôme minimal de $\mathbf{1}_E$.
(2) λ est un polynôme minimal de $\mathbf{0} : E \to E : u \mapsto 0_E$.

8.4.2. Proposition. Soient $T \in \text{End}(E)$ et $A \in M_n(K)$. Si $m(\lambda)$ est un polynôme minimal de T (respectivement, de A), alors $m(\lambda)$ divise tous les annulateurs de T (respectivement, de A) sur K.

Démonstration. Supposons qu'il existe un annulateur $f(\lambda)$ de T non divisible par $m(\lambda)$. Alors $f(\lambda) = m(\lambda)q(\lambda) + r(\lambda)$, où $r(\lambda)$ est non nul de degré plus petit que le degré de $m(\lambda)$. Or

$$r(T) = f(T) - m(T)q(T) = 0,$$

c'est-à-dire, $r(T)$ est un annulateur de T. Ceci contredit la minimalité du degré de $m(\lambda)$. La preuve s'achève.

8.4.3. Corollaire. Tout $T \in \mathrm{End}(E)$ (respectivement, $A \in M_n(K)$) admet un seul polynôme minimal, noté $m_T(\lambda)$ (respectivement, $m_A(\lambda)$).

Démonstration. Soient $m_1(\lambda)$ et $m_2(\lambda)$ des polynômes minimaux de T. Il suit de la proposition 8.4.8 que $m_2(\lambda) \mid m_1(\lambda)$ et $m_1(\lambda) \mid m_2(\lambda)$. Ainsi $m_2(\lambda) = am_1(\lambda)$ avec $a \in K$. En comparant les coefficients, on voit que $a = 1_K$. Ceci achève la démonstration.

8.4.4. Proposition. Soit $T \in \mathrm{End}(E)$. Si A est la matrice de T dans une base de E, alors $m_T(\lambda) = m_A(\lambda)$.

Démonstration. D'après le lemme 8.3.7(2), $m_T(A) = 0$ et $m_A(T) = 0$. D'après la proposition 8.4.2, $m_A(\lambda) \mid m_T(\lambda)$ et $m_T(\lambda) \mid m_A(\lambda)$. Donc $m_A(\lambda) = m_T(\lambda)$ puisque $m_T(\lambda)$ et $m_A(\lambda)$ sont tous normalisés. Ceci achève la démonstration.

Le résultat utile suivant est une autre conséquence immédiate de la proposition 8.4.2.

8.4.5. Corollaire. Soit $A \in M_n(K)$. Si $f(\lambda)$ est un annulateur de A, alors $m_A(\lambda)$ est le facteur normalisé de plus petit degré de $f(\lambda)$ qui annule A.

Exemple. Soit $J_n(\lambda_0)$ un bloc de Jordan. On sait que $(\lambda - \lambda_0)^n$ est un annulateur de $J_n(\lambda_0)$. Donc $m_{J_n(\lambda_0)} = (\lambda - \lambda_0)^r$ avec $1 \le r \le n$. Or $(J_n(\lambda_0) - \lambda_0 I_n)^{n-1} \neq 0$. Donc $(\lambda - \lambda_0)^r$ n'est pas un annulateur de $J_n(\lambda_0)$ pour tout $0 \le r < n$. Par conséquent, $m_{J_n(\lambda_0)} = (\lambda - \lambda_0)^n$.

Dès maintenant, on discutera commet caractériser la diagonalisabilité en terme de polynôme minimal.

8.4.6. Lemme. Soient $T \in \mathrm{End}(E)$ et $A \in M_n(K)$. Les valeurs propres de T (respectivement, de A) sont les racines de $m_T(\lambda)$ (respectivement, de $m_A(\lambda)$) dans K.

Démonstration. D'après le théorème 8.3.8 et la proposition 8.4.2, $m_A(\lambda) | \chi_A(\lambda)$. Ainsi les racines de $m_A(\lambda)$ sont des valeurs propres de A. Réciproquement, si $\lambda_0 \in K$ n'est pas une racine de $m_A(\lambda)$, alors $m_A(\lambda)$ et $\lambda - \lambda_0$ sont co-premiers. Ainsi il existe $p(\lambda), q(\lambda) \in K[\lambda]$ tels que $p(\lambda)m_A(\lambda) + q(\lambda)(\lambda - \lambda_0) = 1$. D'où

$$I_n = q(A)(A - \lambda_0 I_n) = (A - \lambda_0 I_n)q(A),$$

c'est-à-dire, $A - \lambda_0 I_n$ est inversible. D'après la proposition 8.1.9, λ_0 n'est pas une valeur propre de A. La démonstration s'achève.

8.4.7. Théorème. Soit $T \in \mathrm{End}(E)$ dont $f(\lambda)$ est un annulateur. Si $f(\lambda) = f_1(\lambda) \cdots f_r(\lambda)$ avec les $f_i(\lambda)$ deux à deux co-premiers, alors E se décompose comme suit :

$$E = \mathrm{Ker} f_1(T) \oplus \cdots \oplus \mathrm{Ker} f_r(T).$$

Démonstration. Écrivons $f(\lambda) = f_i(\lambda)g_i(\lambda)$, $i = 1, \ldots, r$. Remarquons que le plus grand commun diviseur de $g_1, \cdots, g_r$ est 1. Ainsi il existe $p_1(\lambda), \cdots, p_r(\lambda) \in K[\lambda]$ tels que $g_1p_1 + \cdots + g_rp_r = 1$. D'où,

$g_1(T)p_1(T)+\cdots+g_r(T)p_r(T) = \mathbf{1}_E$. Pour tout $u \in E$, on a $u = \mathbf{1}_E(u) = g_1(T)p_1(T)(u)+\cdots+g_r(T)p_r(T)(u)$. Comme

$$f_i(T)[g_i(T)p_i(T)(u)] = p_i(T)[f(T)(u)] = p(T)(0_E) = 0_E, \ i = 1, \ldots, r,$$

on a $E = \mathrm{Ker} f_1(T) + \cdots + \mathrm{Ker} f_r(T)$. Supposons que $u_1 + \cdots + u_r = 0_E$, où $u_i \in \mathrm{Ker} f_i(T)$, $i = 1, \ldots, r$. Alors $g_i(T)(u_j) = 0_E$ lorsque $i \neq j$. Ainsi $g_i(T)(u_i) = g_i(T)(u_1 + \cdots + u_r) = 0_E$, et donc $u_i = \mathbf{1}_E(u_i) = g_i(T)p_i(T)(u_i) = p_i(T)g_i(T)(u_i) = 0_E$, $i = 1, \ldots, r$. Ceci montre que $E = \mathrm{Ker} f_1(T) \oplus \cdots \oplus \mathrm{Ker} f_r(T)$. La preuve s'achève.

8.4.8. Théorème. Une matrice $A \in M_n(K)$ est diagonalisable sur K si et seulement si $m_A(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_r)$ avec $\lambda_1, \ldots, \lambda_r \in K$ deux à deux distincts. Dans ce cas, $\lambda_1, \ldots, \lambda_r \in K$ sont les valeurs propres distinctes de A.

Démonstration. Considérons l'endomorphisme de K-espace vectoriel $K^{(n)}$ suivant :

$$T : K^{(n)} \to K^{(n)} : u \mapsto uA.$$

Alors $m_T(\lambda) = m_A(\lambda)$ et T est diagonalisable si et selement si A est diagonalisable sur K.

Supposons que T est diagonalisable. Alors $K^{(n)}$ a une base $\{u_1, \ldots, u_n\}$ telle que $T(u_i) = \lambda_i u_i$, $\lambda_i \in K$, $i = 1, \ldots, n$. On suppose que $\lambda_1, \ldots, \lambda_r$ avec $r \leq n$ sont deux à deux distincts tels que $\{\lambda_1, \ldots, \lambda_r\} = \{\lambda_1, \ldots, \lambda_n\}$. Posons $m(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_r)$. D'après le lemme 8.4.6, $m(\lambda) \mid m_T(\lambda)$. D'autre part, on se fixe un indice quelconque i avec $1 \leq i \leq n$. Remarquons $T(u_i) = \lambda_j u_i$, c'est-à-dire, $(T - \lambda_j \mathbf{1})(u_i) = 0_E$, pour un certain $1 \leq j \leq r$. D'où $m(T)(u_i) = 0_E$. Comme $\{u_1, \ldots, u_n\}$ est une base, $m(T) = 0$. Donc $m_T(\lambda) \mid m(\lambda)$. Ainsi $m_T(\lambda) = m(\lambda)$.

Réciproquement supposons que $m_T(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_r)$, où $\lambda_1, \ldots, \lambda_r \in K$ sont deux à deux distincts. Alors $\lambda - \lambda_1, \ldots, \lambda - \lambda_r$ sont deux à deux co-premiers. Donc

$$K^{(n)} = \mathrm{Ker}(T - \lambda_1 \mathbf{1}) \oplus \cdots \oplus \mathrm{Ker}(T - \lambda_r \mathbf{1}).$$

Remarquons que $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres de T. Donc $\mathrm{Ker}(T - \lambda_i \mathbf{1}) \neq 0$, pour tout $1 \leq i \leq r$. Prenons une base $\mathcal{B}_i$ de $\mathrm{Ker}(T - \lambda_i \mathbf{1})$. Alors $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est une base de $K^{(n)}$, ce qui est formée de vecteurs propres de T. Donc T est diagonalisable. Ceci achève la démonstration.

Exemple. Soit $J_n(\lambda_0)$ un bloc de Jordan. On sait que $m_{J_n(\lambda_0)} = (\lambda - \lambda_0)^n$. Donc $J_n(\lambda_0)$ est diagonalisable si et seulement si $n = 1$.

Le résultat suivant est une conséquence immédiate du théorème précédent.

8.4.9. Corollaire. Si $A \in M_n(K))$ admet un annulateur $f(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_r)$ avec $\lambda_1, \ldots, \lambda_r \in K$ deux à deux distincts, alors A est diagonalisable.

8.5. Formes de Jordan

Partout dans cette section, on se fixe T un endomorphisme de E.

8.5.1. Définition. Un sous-espace F de E est dit *stable* par T (ou brièvement, T-*stable*) si $T(F) \subseteq F$ (c'est-à-dire, $T(u) \in F$ pour tout $u \in F$). Dans ce cas, $T|_F : F \to F : u \mapsto T(u)$ est un endomorphisme de F, noté encore T.

Exemple. (1) E et $\{0_E\}$ sont T-stables.
(2) Soit $u \in E$ non nul. Alors u est un vecteur propre de T si et seulement si $< u >$ est T-stable.
(3) Soit $f(\lambda) \in K[\lambda]$. Alors Ker$f(T)$ est T-stable. En effet, pour tout $u \in$ Ker$f(T)$,

$$f(T)(T(u)) = (f(T)T)(u) = (Tf(T))(u) = T(f(T)(u)) = T(0_E) = 0_E.$$

Donc $T(u) \in$Ker$f(T)$.

(4) Considérons

$$T : \mathbb{R}^4 \to \mathbb{R}^4 : (x_1, x_2, x_3, x_4) \mapsto (2x_1 - x_2, x_1 + x_2, 5x_3 - 4x_4, 3x_3 + 3x_4).$$

Alors $F_1 = \{(x_1, x_2, 0, 0) \mid x_1, x_2 \in \mathbb{R}\}$ est T-stable. En effet, si $u = (x_1, x_2, 0, 0) \in F_1$, alors $T(u) = (2x_1 - x_2, x_1 + x_2, 0, 0) \in F_1$.

De même façon, on voit que $F_2 = \{(0, 0, x_3, x_4) \mid x_3, x_4 \in \mathbb{R}\}$ est T-stable; mais que $F_3 = (0, x_2, x_3, 0) \mid x_2, x_3 \in \mathbb{R}\}$ n'est pas T-stable.

8.5.2. Lemme. Soit F un sous-espace de E ayant pour base $\{v_1, \ldots, v_r\}$. Alors F est T-stable si et seulement si $T(v_i) \in F$, pour tout $1 \le i \le r$.

Démonstration. Supposons que $T(v_i) \in F, i = 1, \ldots, r$. Or tout $u \in F$ s'écrit comme $u = \alpha_1 v_1 + \cdots + \alpha_r v_r, \alpha_i \in K$. Ainsi $T(u) = \alpha_1 T(v_1) + \cdots + \alpha_r T(v_r) \in F$. Par conséquent, F est T-stable.

8.5.3. Proposition. Soit $E = E_1 \oplus E_2$ avec $E_1,\ E_2$ non nuls et T-stables. Soit $\mathcal{B}_i$ une base de E_i dans la quelle la matrice de T est $A_i,\ i = 1, 2$. Alors $\mathcal{B}_1 \cup \mathcal{B}_2$ est une base de E dans laquelle la matrice de T est

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

Démonstration. Posons $\mathcal{B}_1 = \{u_1, \ldots, u_r\}$, $\mathcal{B}_2 = \{u_{r+1}, \ldots, u_n\}$, $A_1 = (a_{ij})_{1 \le i,j \le r}$ et $A_2 = (a_{ij})_{r+1 \le i,j \le n}$. Alors

$$\begin{array}{rcrcccrcrcccr}
T(u_1) & = & a_{11}u_1 & + & \cdots & + & a_{r1}u_r & + & 0u_{r+1} & + & \cdots & + & 0u_n \\
\vdots & & \vdots & & & & \vdots & & \vdots & & & & \vdots \\
T(u_r) & = & a_{1r}u_1 & + & \cdots & + & a_{rr}u_r & + & 0u_{r+1} & + & \cdots & + & 0u_n \\
T(u_{r+1}) & = & 0u_1 & + & \cdots & + & 0u_r & + & a_{r+1,r+1}u_{r+1} & + & \cdots & + & a_{n,r+1}u_n \\
\vdots & & \vdots & & & & \vdots & & \vdots & & & & \vdots \\
T(u_n) & = & 0u_1 & + & \cdots & + & 0u_r & + & \alpha_{r+1,n}u_{r+1} & + & \cdots & + & \alpha_{n,n}u_n
\end{array}$$

Par conséquent, la matrice de T dans $\mathcal{B}_1 \cup \mathcal{B}_2$ est

$$A = \begin{pmatrix}
a_{11} & \cdots & a_{1r} & 0 & \cdots & 0 \\
\vdots & & \vdots & \vdots & & \vdots \\
a_{r1} & \cdots & a_{rr} & 0 & \cdots & 0 \\
0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{r+1,n} \\
\vdots & & \vdots & \vdots & & \vdots \\
0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{n,n}
\end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

Exemple. Considérons

$$T : \mathbb{R}^4 \to \mathbb{R}^4 : (x_1, x_2, x_3, x_4) \mapsto (2x_1 - x_2, x_1 + x_2, 5x_3 - 4x_4, 3x_3 + 3x_4)$$

Alors $F_1 =< e_1, e_2 >$ et $F_2 =< e_3, e_4 >$ sont T-stables tels que $\mathbb{R}^4 = F_1 \oplus F_2$. Or $T(e_1) = 2e_1 + e_2,\ T(e_2) = -e_1 + e_2$. Donc la matrice de T dans $\{e_1, e_2\}$ est

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}.$$

De même la matrice de T dans $\{e_3, e_4\}$ est

$$\begin{pmatrix} 5 & -4 \\ 3 & 3 \end{pmatrix}.$$

Ainsi la matrice de T dans $\{e_1, e_2, e_3, e_4\}$ est

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 5 & -4 \\ 0 & 0 & 3 & 3 \end{pmatrix}.$$

8.5.4. Corollaire. Soit

$$E = E_1 \oplus E_2 \oplus \cdots \oplus E_r$$

avec les E_i non nuls T-stables. Soit $\mathcal{B}_i$ une base de E_i dans laquelle la matrice de T est $A_i,\ i = 1, \ldots, r$. Alors $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$ est une base de E dans laquelle la matrice de T est

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}.$$

8.5.5. Corollaire. T est diagonalisable si et seulement si E se décompose en somme directe de sous-espaces T-stables de dimension 1.

Démonstration. T est diagonalisable si et seulement si il existe une base $\{u_1, \ldots, u_n\}$ de E telle que $T(u_i) = \lambda_i u_i,\ \lambda_i \in K,\ i = 1, \ldots, n$ si et seulement $< u_1 >, \ldots, < u_r >$ sont T-stables de dimension 1 tels que

$$E =< u_1 > \oplus \cdots \oplus < u_n > .$$

8.5.6. Définition. Soit F un sous-espace non nul et T-stable de E. On dit que F est T-*décomposable* si $F = F_1 \oplus F_2$ avec F_1, F_2 non nuls T-stables; et T-*indécomposable* sinon.

Exemple. (1) Si $\dim E = 1$, alors E est T-indécomposable.

(2) Considérons $\mathbb{R}^4$ et

$$T : \mathbb{R}^4 \to \mathbb{R}^4 : (x_1, x_2, x_3, x_4) \mapsto (2x_1 - x_2, x_1 + x_2, 5x_3 - 4x_4, 3x_3 + 3x_4).$$

On sait que $F_1 =< e_1, e_2 >$ et $F_2 =< e_3, e_4 >$ sont T-stables tels que $\mathbb{R}^4 = F_1 \oplus F_2$. Donc $\mathbb{R}^4$ est T-décomposable.

8.5.7. Proposition. Soit $\dim(E) = n > 0$. Alors E est T-décomposable si et seulement s'il existe une base $\{u_1, \ldots, u_n\}$ de E dans la quelle la matrice de T est de la forme

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

avec A_1, A_2 des matrices carrées d'ordre $< n$.

Démonstration. La nécessité suit de la proposition 8.5.3.

Pour prouver la suffisance, posons $A_1 = (a_{ij})_{1 \le i,j \le r}$ et $A_2 = (a_{ij})_{r+1 \le i,j \le n}$ avec $1 \le r < n$. Posons

$$F_1 =< u_1, \ldots, u_r > \quad \text{et} \quad F_2 =< u_{r+1}, \ldots, u_n > .$$

Alors $E = F_1 \oplus F_2$ avec F_1, F_2 non nuls. Pour tout $1 \le j \le r,\ \ T(u_j) = \sum_{i=1}^{r} a_{ij} u_i \in F_1$; et pour tout $r+1 \le j \le n,\ \ T(u_j) = \sum_{i=r+1}^{n} a_{ij} u_i \in F_2$. Donc F_1 et F_2 sont T-stables. Ainsi E est T-décomposable.

8.5.8. Théorème. Soit E non nul. Alors E se décompose

$$E = E_1 \oplus \cdots \oplus E_r,$$

où $E_1, \ldots, E_r$ sont T-stables et T-indécomposables.

Démonstration. Procédons par récurrence en n =dim(E). Si $n = 1$, alors E lui-même est T-indécomposable. Supposons que $n > 1$ et le résultat est valide pour les espaces de dimension $< n$. Si E est T-indécomposable, alors il n'y a rien à prouver. Sinon $E = F_1 \oplus F_2$ avec F_1, F_2 non nuls et T-stables. Donc $\dim(F_i) < n,\ i = 1, 2$. Par hypothèse de récurrence $F_1 = E_1 \oplus \cdots \oplus E_s$ et $F_2 = E_{s+1} \oplus \cdots \oplus E_r$ avec les E_i étant T-stables et T-indécomposables. Or

$$E = E_1 \oplus \cdots \oplus E_s \oplus E_{s+1} \oplus \cdots \oplus E_r$$

est une décomposition désirée.

Remarque. D'après le corollaire 8.4.4, plus $\dim(E_1), \ldots, \dim(E_r)$ sont petites, plus T est près d'être diagonalisable.

8.5.9. Définition. On dit que T est *nilpotent* d'indice de nilpotence r si $T^r = 0$ et T^{r-1} est non nul.

Remarque. T est nilpotent d'indice de nilpotence r si et seulement si sa matrice dans une base est nilpotente d'indice de nilpotence r.

Exemple. (1) Considérons

$$J_n(0) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

On sait que $(J_n(0))^n = 0$ et

$$(J_n(0))^{n-1} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Ainsi $J_n(0)$ est nilpotente d'indice de nilpotence n.
(2) Considérons $T : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \mapsto (x, y, z)A$, où

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Alors la matrice de T dans la base canonique de $\mathbb{R}^3$ est $A^T = J_3(0)$, ce qui est nilpotente d'indice de nilpotence 3. Ainsi T est nilpotent d'indice de nilpotence 3.

8.5.10. Lemme. Soit T nilpotent d'indice de nilpotence r. Alors il existe $u \in E$ tel que $\{u, T(u), \ldots, T^{r-1}(u)\}$ est libre. Par conséquent, $r \leq \dim(E)$.

Démonstration. Comme $T^{r-1} \neq 0$, il existe $u \in E$ tel que $T^{r-1}(u) \neq 0_E$. On prétend que $\{T^{r-s}(u), \ldots, T^{r-1}(u)\}$ est libre pour tout $1 \leq s \leq r$. D'abord, $\{T^{r-1}(u)\}$ est libre. Supposons que $1 \leq s < r$ et $\{T^{r-s)}(u), \ldots, T^{r-1}(u)\}$ est libre. Supposons que

$$\alpha_{r-(s+1)}T^{r-(s+1)}(u) + \alpha_{r-s}T^{r-s}(u) + \cdots + \alpha_{r-2}T^{r-2}(u) + \alpha_{r-1}T^{r-1}(u) = 0_E, \ \alpha_i \in K.$$

En appliquant T, on obtient

$$\alpha_{r-(s+1)}T^{r-s}(u) + \cdots + \alpha_{r-2}T^{r-2}(u) = 0_E$$

car $T^r = 0$. Donc $\alpha_{r-(s+1)} = \cdots = \alpha_{r-2} = 0$. Ainsi $\alpha_{r-1}T^{r-1}(u) = 0_E$, et donc $\alpha_{r-1} = 0$ car $T^{r-1}(u)$ est non nul. Ce qui montre que $\{u, T(u), \ldots, T^{r-1}(u)\}$ est libre. Ceci achève la démonstration.

8.5.11. Théorème. Soit E de dimension n. Supposons que T est nilpotent d'indice de nilpotence r. Les conditions suivantes sont équivalentes:
(1) E est T-indécomposable.
(2) $r = n$.
(3) Il existe une base de E dans laquelle la matrice de T est $J_n(0)$.

Démonstration. On accepte sans preuve le fait que (1) implique (2).

Supposons maintenant que $r = n$. Alors il existe $u \in E$ tel que $\mathcal{B} = \{u, T(u), \ldots, T^{n-1}(u)\}$ est libre, et donc une base de E. Or

$$\big(T(u), T(Tu), \ldots, T(T^{n-1}u)\big) = \big(u, T(u), \ldots, T^{n-1}(u)\big) \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

Par conséquent la matrice de T dans $\mathcal{B}$ est $J_n(0)$.

Supposons enfin que $J_n(0)$ est la matrice de T dans une base de E. Comme $J_n(0)$ est d'indice de nilpotence n, on a $r = n$. Si E est T-décomposable, alors $E = F_1 \oplus F_2$ avec F_1, F_2 non nuls et T-stables. Posons $T_i = T|_{F_i}$. Alors T_i est nilpotent, disons d'indice de nilpotence r_i. D'après le lemme 8.5.10, $r_i \leq \dim(F_i) < n$, $i = 1, 2$. Prenons $s = \max\{r_1, r_2\}$. Alors $s < n$ et $T_1^s = T_2^s = 0$. Or tout $u \in E$ s'écrit $u = u_1 + u_2$, $u_i \in F_i$. Ainsi $T^s(u) = T^s(u_1) + T^s(u_2) = T_1^s(u_1) + T_2^s(u_2) = 0_E + 0_E = 0_E$. Ainsi $T^s = 0$, et donc $T^{n-1} = 0$ car $s \leq n - 1$, Ce qui contredit que T est d'indice de nilpotence n. Donc E est T-indécomposable.

Exemple. Considérons

$$T : \mathbb{R}^{(3)} \to \mathbb{R}^{(3)} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

où

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

On sait que A est la matrice de T dans la base canonique de $\mathbb{R}^{(3)}$. Or A est nilpotente d'indice de nilpotence 3. Ainsi T est nilpotent d'indice de nilpotence 3. D'après le théorème 8.5.11, il existe une base de $\mathbb{R}^{(3)}$ dans la quelle la matrice de T est $J_3(0)$. Par conséquent, A est semblable à $J_3(0)$.

8.5.12. Corollaire. Soit T nilpotent. Alors il existe une base $\mathcal{B}$ de E telle que

$$[T]_{\mathcal{B}} = \begin{pmatrix} J_{n_1}(0) & 0 & \cdots & 0 \\ 0 & J_{n_2}(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n_r}(0) \end{pmatrix}.$$

Démonstration. D'après le théorme 8.4.8, $E = E_1 \oplus \cdots \oplus E_r$, où $E_1, \ldots, E_r$ sont T-stables et T-indécomposables. D'après le théorème 8.4.11, pour tout $1 \leq i \leq r$, il existe une base $\mathcal{B}_i$ de E_i dans laquelle la matrice de T est $J_{n_i}(0)$. Ainsi $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est une base de E dans laquelle la matrice de T est

$$\begin{pmatrix} J_{n_1}(0) & 0 & \cdots & 0 \\ 0 & J_{n_2}(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n_r}(0) \end{pmatrix}.$$

8.5.13. Définition. Une matrice carrée A est dite *sous la forme de Jordan* si

$$A = \begin{pmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n_r}(\lambda_r) \end{pmatrix}.$$

Remarque. Si

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_s \end{pmatrix}$$

où les A_i sont toutes sous la forme de Jordan, alors A l'est aussi.

Exemple. (1) Une forme de Jordan est diagonale si et seulement si tous les blocs de Jordan sont d'ordre 1.

(2) La matrice

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

est sous la forme de Jordan.

(3) La matrice

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

n'est pas sous la forme de Jordan.

8.5.14. Théorème. Si $\chi_T(\lambda)$ est scindé sur K, alors il existe une base $\mathcal{B}$ de E dans laquelle la matrice de T est sous la forme canonique de Jordan, ce qui est unique à ordre près des blocs de Jordan.

Démonstration. Supposons que

$$\chi_T(\lambda) = (-1)^n (\lambda - \lambda_1)^{n_i} \cdots (\lambda - \lambda_r)^{n_r},$$

où $\lambda_1, \ldots, \lambda_r \in K$ sont deux à deux distincts et $n_i > 0$. D'après le théorème de Hamilton-Cayley, $f(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_r)^{n_r}$ est un annulateur de T avec $(\lambda - \lambda_1)^{n_1}, \ldots, (\lambda - \lambda_r)^{n_r}$ sont deux à deux co-premières. Posons $E_i = \text{Ker}(T - \lambda_i \mathbf{1})^{n_i}$ ce qui est T-stable. Alors

$$E = E_1 \oplus \cdots \oplus E_r.$$

Or pour tout $u \in E_i$, $(T - \lambda_i \mathbf{1})^{n_i}(u) = 0$. Donc $T - \lambda_i \mathbf{1}$ est nilpotent sur E_i. D'après le corollaire 8.4.12, il existe une base $\mathcal{B}_i$ de E_i telle que

$$[T - \lambda_i \mathbf{1}]_{\mathcal{B}_i} = \begin{pmatrix} J_{m_1}(0) & 0 & \cdots & 0 \\ 0 & J_{m_2}(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_s}(0) \end{pmatrix}_{n_i \times n_i}.$$

Par conséquent,

$$[T]_{\mathcal{B}_i} = [(T - \lambda_i \mathbf{1}) + \lambda_i \mathbf{1}]_{\mathcal{B}_i} = [T - \lambda_i \mathbf{1}]_{\mathcal{B}_i} + [\lambda_i \mathbf{1}]_{\mathcal{B}_i} = \begin{pmatrix} J_{m_1}(\lambda_i) & 0 & \cdots & 0 \\ 0 & J_{m_2}(\lambda_i) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_s}(\lambda_i) \end{pmatrix}$$

est sous la forme de Jordan. D'après le corollaire 8.5.4, $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est une base de E telle que

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T]_{\mathcal{B}_1} & 0 & \cdots & 0 \\ 0 & [T]_{\mathcal{B}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [T]_{\mathcal{B}_r} \end{pmatrix}.$$

Comme $\mathbb{C}$ est algébriquement clos, le résultat suivant est une conséqence immédiate du théorème précédent.

8.5.15. Théorème. Soit E un espace complexe de dimension finie. Alors pour tout $T \in \operatorname{End}(E)$, il existe une base $\mathcal{B}$ de E telle que $[T]_{\mathcal{B}}$ est sous la forme de Jordan, ce qui est unique à ordre près des blocs de Jordan.

8.5.16. Théorème. Toute matrice carrée complexe est semblable à une forme de Jordan, ce qui est unique à ordre près des blocs de Jordan.

8.5.17. Calcul de la forme de Jordan. Soit $A \in M_n(K)$ et supposons que $\chi_A(\lambda) = (\lambda_1 - \lambda)^{n_1} \cdots (\lambda_r - \lambda)^{n_r}$, où $\lambda_1, \ldots, \lambda_r$ sont deux à deux distincts et $n_i > 0$. Une forme de Jordan de A est alors constituée des blocs de Jordan correspondants aux valeurs propres $\lambda_1, \ldots, \lambda_r$. Pour chaque $1 \le i \le r$, on peut trouver les blocs de Jordan correspondant à λ_i comme suit:

(1) Calculer r_j =rg$(A - \lambda_i I)^j, j = 0, 1, 2, \cdots,$ jusqu'à trouver un premier nombre m_i tel que rg$(A - \lambda_i I)^{m_i}$ =rg$(A - \lambda_i I)^{m_i+1}$.

(2) L'ordre maximal des blocs de Jordan correspondants à λ_i est m_i. Et pour tout $1 \le j \le m_i$, le nombre des blocs de Jordan d'ordre j correspondants à λ_i est $s_j(\lambda_i) = (r_{j-1} + r_{j+1}) - 2r_j$.

Exemple. (1) Trouver une forme de Jordan de la matrice

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & 3 & 0 \\ 1 & 0 & 1 & 1 & 3 \end{pmatrix}.$$

Solution. $\chi_A(\lambda) = (2 - \lambda)^3 (3 - \lambda)^2$. Les valeurs propres distinctes sont $\lambda_1 = 2$ et $\lambda_2 = 3$.

(a) Blocs de Jordan corréspondants à λ_1.

r_0 =rg$(A - 2I)^0 = 5$, $r_1 =$ rg$(A - 2I) = 3$, r_2 =rg$(A - 2I)^2 = 2$ et r_3 =rg$(A - 2I)^3 = 2$. Donc l'ordre maximal des blocs de Jordan correspondant à λ_1 est 2. On a que $s_1(\lambda_1) = r_0 + r_2 - 2r_1 = 7 - 6 = 1$ et $s_2(\lambda_1) = (r_1 + r_3) - 2r_2 = 5 - 4 = 1$. Donc les blocs de Jordan correspondants à λ_1 sont $J_1(2)$ et $J_2(2)$.

(b) Blocs de Jordan correspondants à λ_2.

r_0 =rg$(A - 3I)^0 = 5$, r_1 =rg$(A - 3I) = 4$, r_2 =rg$(A - 2I)^2 = 3$ et r_3 =rg$(A - 2I)^3 = 3$. Donc l'ordre maximal des blocs de Jordan correspondants à λ_2 est 2. On a que $s_1(\lambda_2) = (r_2 + r_0) - 2r_1 = 8 - 8 = 0$ et $s_2(\lambda_2) = (r_3 + r_1) = 7 - 6 = 1$. Donc le bloc de Jordan correspondant à λ_2 est $J_2(3)$.

Donc les blocs de Jordan de A sont $J_1(2), J_2(2)$ et $J_2(3)$. Ainsi on trouve une forme de Jordan de A comme suit:

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}.$$

(2) Trouver une forme de Jordan de

$$A = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 1 \\ 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}.$$

Solution. $\chi_A(\lambda) = (5-\lambda)^5$. Donc $\lambda_1 = 5$ est la seule valeur propre distincte de A.

Or $r_0 = \text{rg}(A-5I)^0 = 5$, $r_1 = \text{rg}(A-5I) = 3$, $r_2 = \text{rg}(A-5I)^2 = 2$, $r_3 = \text{rg}(A-5I)^3 = 1$, $r_4 = \text{rg}(A-5I)^4 = 0$ et $r_5 = \text{rg}(A-5I)^5 = 0$. Donc l'ordre maximal des blocs de Jordan correspondant à λ_1 est 4. On a que $s_1 = (r_2 + r_0) - 2r_1 = 7 - 6 = 1$ et $s_2 = (r_3 + r_1) - 2r_2 = 4 - 4 = 0, s_3 = (r_4 + r_2) - 2r_3 = 2 - 2 = 0$ et $s_4 = (r_5 + r_3) - 2r_4 = 1 - 0 = 1$. Donc les blocs de Jordan corrépondants sont $J_1(5)$ et $J_4(5)$. Donc on trouve une forme de Jordan de A comme suit:

$$\begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 & 5 \end{pmatrix}.$$

8.6. Exercices

1. Déterminer quelle valeur de 2 et -2 est une valeur propre de la matrice suivante:

$$\begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}.$$

2. Soit

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Vérifier que 3 et 2 sont des valeurs propres de A, et trouver leur multiplicité géométrique.

3. Considérer la matrice suivante:
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 1 & \alpha \end{pmatrix}.$$
Pour quelle valeur de α, la matrice A, admet-t-elle la valeur 2 comme valeur propre? Dans ce cas, trouver une base de l'espace propre de A associé à 2.

4. Trouver les valeurs propres et leur multiplicité géométrique de chacune des matrices suivantes:
$$(1) \quad \begin{pmatrix} 7 & -6 & -9 \\ -2 & 6 & -6 \\ -1 & -2 & 7 \end{pmatrix}; \qquad (2) \quad \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

5. (1) Si $A \in M_n(K)$, montrer que A et A^T ont les mêmes valeurs propres.
   (2) Donner une matrice A telle que A et A^T n'aient pas les mêmes vecteurs propres.

6. Montrer qu'une matrice carrée A est non inversible si et seulement si la valeur 0 est une valeur propre de A.

7. Soit A une matrice inversible. Montrer que λ_0 est une valeur propre de A si et seulement si $\dfrac{1}{\lambda_0}$ est une valeur propre de A^{-1}.

8. Soient A et B deux matrices carrées semblables avec $B = P^{-1}AP$. Comparer les valeurs propres et les vecteurs propres de A avec ceux de B.

9. Soient $A \in M_n(K)$ et $\alpha \in K$. Comparer les valeurs propres et les vecteurs propres de A avec ceux de $A + \alpha I_n$.

10. Montrer qu'une matrice carrée complexe admet toujours une vecteur propre dans $\mathbb{C}$.

11. Soit A une matrice carrée réelle. Si λ_0 est une valeur propre complexe de A, montrer que $\overline{\lambda_0}$ l'est également.

12. Soit E un espace vectoriel réel de dimension finie impaire. Montrer que tout endomorphisme de E admet au moins un vecteur propre.

13. Soit A une matrice carrée telle que $\chi_A(\lambda)$ est scindé sur K. Montrer que si $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres deux à deux distinctes de A, alors
$$\lambda_1^{n_1} \cdots \lambda_r^{n_r} = \det(A) \quad \text{et} \quad n_1\lambda_1 + \cdots + n_r\lambda_r = \operatorname{tr}(A),$$
où n_i est la multiplicité algébrique de λ_i pour tout $1 \leq i \leq r$.

14. Pour quelles valeurs de $a, b, c \in \mathbb{R}$, la matrice suivante, est-elle diagonalisable sur $\mathbb{R}$?
$$\begin{pmatrix} 1 & a & 1 \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix}.$$

15. Pour quelles valeurs de $a \in \mathbb{Q}$, la matrice suivante, est-elle diagonalisable sur $\mathbb{Q}$?

$$\begin{pmatrix} 1 & a & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

16. Montrer que la matrice complexe suivate est diagonalisable:

$$A = \begin{pmatrix} a & b & \cdots & b & b \\ b & a & \cdots & b & b \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & \cdots & b & a \end{pmatrix}_{n\times n}.$$

17. Soit $A = uv^T$, où $u, v \in K^{(n)}$ avec $n > 1$. Montrer les énoncés suivants:
    (1) Les valeurs propres de A sont 0 et $v^T u$.
    (2) A est diagonalisable si et seulement si l'un de u, v est nul ou bien $v^T u \neq 0$.

    Indication: Considérer séparément les cas où l'un des u et v est nul et où aucun de u et v n'est nul.

18. Soit A une matrice carrée diagonalisable. Montrer que $\text{rg}(A^r) = \text{rg}(A)$ pour tout $r > 0$.

19. Diagonaliser si possible chacune des matrices suivantes:

$$(1) \quad \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}; \qquad (2) \quad \begin{pmatrix} -4 & -4 & -8 \\ 4 & 6 & 4 \\ 6 & 4 & 10 \end{pmatrix}.$$

20. Diagonaliser si possible l'endomorphisme $T : \mathbb{C}^{(2)} \to \mathbb{C}^{(2)} : au_1 + bu_2 \mapsto -bu_1 + au_2$, où

$$\left\{ u_1 = \begin{pmatrix} i \\ i \end{pmatrix}, u_2 = \begin{pmatrix} 1+i \\ 1-i \end{pmatrix} \right\}.$$

21. Diagonaliser l'endomorphisme T de $\mathbb{R}_4[x]$ dont la matrice dans la base canonique est la matrice suivante:

$$A = \begin{pmatrix} 5 & 1 & 1 & -1 \\ 1 & 5 & 1 & -1 \\ 1 & 1 & 5 & -1 \\ -1 & -1 & -1 & 5 \end{pmatrix}.$$

22. Calculer $I + 3A + A^{150}$ pour

$$A = \begin{pmatrix} 6 & -3 \\ 4 & -1 \end{pmatrix}.$$

23. Calculer $I_2 + A^{100}$, où

$$A = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}.$$

24. Soit $f(\lambda) \in K[\lambda]$. Montrer que si λ_0 est une valeur propre d'une matrice carrée A, alors $f(\lambda_0)$ est une valeur propre de $f(A)$.

25. (1) Soit $A \in M_n(K)$. Montrer que pour tout $f(\lambda) \in K[\lambda]$, il existe $r(\lambda) \in K[\lambda]$ de degré $< n$ tel que $f(A) = r(A)$.

    (2) Considérons la matrice rationnelle

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

    Calculer $A^6 - A^4 + A^2 - 4I_3$ en utilisant la partie (1) et le Théorème de Hamilton-Cayley.

26. Trouver le polynôme minimal de chacune des matrices suivantes et, si possible, calculer l'inverse au moyen de celui-ci.

$$(1) \quad \begin{pmatrix} -7 & 0 \\ 0 & -7 \end{pmatrix}; \qquad (2) \quad \begin{pmatrix} 2 & 0 & -2 \\ 0 & 5 & 4 \\ 1 & 0 & -1 \end{pmatrix}; \qquad (3) \quad \begin{pmatrix} -2 & 0 & -2 \\ 5 & 3 & 2 \\ 0 & 3 & 3 \end{pmatrix}.$$

27. Trouver le polynôme minimal de l'endomorphisme de l'espace réel $\mathbb{R}_4[x]$ suivant:

$$T : \mathbb{R}_4[x] \to \mathbb{R}_4[x] : a + bx + cx^2 + dx^3 \mapsto (2a + b) + (2b)x + (3c + d)x^2 + 3dx^3.$$

28. Soit $A \in M_n(K)$. Si A est nilpotente, montrer que
    (2) $A^r = 0$ pour un certain r avec $1 \le r \le n$,
    (1) la valeur 0 est une et la seule valeur propre de A, et
    (3) A est diagonalisable si et seulement si A est nulle.

29. Soit $A \in M_n(K)$. Montrer que A est inversible si et seulement si le terme constant de $m_A(\lambda)$ est non nul.

30. Trouver le polynôme minimal de A et en déduire que A est diagonalisable ou non, où

$$A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}.$$

31. Soit

$$A = \begin{pmatrix} 5 & 1 & 1 & -1 \\ 1 & 5 & 1 & -1 \\ 1 & 1 & 5 & -1 \\ -1 & -1 & -1 & 5 \end{pmatrix}.$$

    Montrer que A est diagonalisable et inversible. En déduire le polynôme minimal de A et calculer A^{-1}.

32. Montrer qu'une matrice idempotente A (c'est-à-dire, $A^2 = A$) est diagonalisable.

33. Soit A une matrice carrée complexe. Montrer les énoncés suivants.
    (1) A est nilpotente si et seulement si la valeur 0 est la seulement valeur propre de A.
    (2) A est diagonalisable si $A^r = A$ pour un entier $r > 1$. *Indication*: Considérer les racines $(r-1)$-ième de l'unité.

34. Considérer $T : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \mapsto (y, -x, z)$. Montrer que $\mathbb{R}^3$ est T-décomposable mais que T n'est pas diagonalisable.

35. Considérer l'endomorphisme de l'espace rationnel $\mathbb{Q}^{(3)}$ suivant:

$$T : \mathbb{Q}^{(3)} \to \mathbb{Q}^{(3)} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & -1 \\ -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

(1) Déterminer lesquels des sous-espaces suivants de $\mathbb{Q}^{(3)}$ sont stables pour T:

$$F_1 = < \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} >; \qquad F_2 = < \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} >.$$

(2) Déterminer si $\mathbb{Q}^{(3)}$ est T-décomposable ou non.

36. Soit $A = (a_{ij})_{n\times n}$ une matrice triangulaire (supérieure ou inférieure) sur un corps K. Montrer que si les termes diagonaux de A sont deux à deux distincts, alors A est semblable à la matrice diagonale $\mathrm{diag}\{a_{11}, \ldots, a_{nn}\}$.

37. Dans chacun des cas suivants, trouver une forme de Jordan de la matrice.

$$(1) \begin{pmatrix} 4 & 0 & 0 \\ 2 & 1 & 3 \\ 5 & 0 & 4 \end{pmatrix}; \qquad (2) \begin{pmatrix} -3 & 0 & 1 \\ 2 & -2 & 1 \\ -1 & 0 & -1 \end{pmatrix}; \qquad (3) \begin{pmatrix} i & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 2 & 0 & -1 & 0 & 2 \end{pmatrix}.$$

38. Trouver une forme de Jordan de

$$A = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 1 \\ 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}.$$

39. Montrer qu'une forme de Jordan J est diagonalisable si et seulement si J est diagonale.

40. Soit A une matrice carrée d'ordre n sur un corps K. Montrer que A est diagonalisable sur K si, et seulement si, les conditions suivantes sont satisfaites:
(1) $\chi_A(\lambda)$ est scindé sur K.
(2) Pour toute valeur propre λ_0 de$A$, $\mathrm{rg}(A - \lambda_0 I_n) = \mathrm{rg}(A - \lambda_0 I_n)^2$.

41. Une matrice carrée A d'ordre n est dit *décomposable* si A est semblable à une matrice de la forme

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

avec A_1 et A_2 des matrices carrées d'ordre $< n$; et *indécomposable* sinon. Montrer qu'une matrice carrée complexe est indécomposable si et seulement si A est semblable à un bloc de Jordan.

Chapitre IX: Espaces euclidiens

Dans le plan, outre la structure algébrique d'espace vectoriel, il y a des autres structures géométriques comme la longueur d'un vecteur, la distance et l'angle entre deux vecteurs. On généralisera ces notions aux espaces réels arbitraires.

9.1. Formes bilinéaires

Le but de cette section est d'étudier les formes bilinéaires sur un espace général. Partout dans cette section, on se fixe E un espace vectoriel sur un corps K.

9.1.1. Définition. Une *forme bilinéaire* sur E est une fonction

$$f : E \times E \to K : (u, v) \mapsto f(u, v)$$

telle que pour tous $u, v, w \in E$ et $\alpha, \beta \in K$,

(1) $f(\alpha u + \beta v, w) = \alpha f(u, w) + \beta f(v, w)$;

(2) $f(w, \alpha u + \beta v) = \alpha f(w, u) + \beta f(w, v)$.

En outre, on dit que f est *symétrique* si $f(u, v) = f(v, u)$, pour tous $u, v \in E$.

Remarque. Pour tous $\alpha, \beta \in K$ et $u, v \in E$, on a $f(\alpha u, v) = \alpha f(u, v)$ et $f(u, \beta v) = \beta f(u, v)$. En particulier, $f(0_E, u) = f(u, 0_E) = 0_K$.

Exemple. Soient $a, b \in \mathbb{R}$ avec $a < b$. Considérons l'espace réel $C[a, b]$ des fonctions continues $f : [a, b] \to \mathbb{R}$. Pour $f, g \in C[a, b]$, on pose

$$< f, g >= \int_a^b f(t)g(t)dt.$$

Ceci donne une forme bilinéaire symétrique sur $C[a, b]$.

9.1.2. Lemme. Si $A = (a_{ij}) \in M_n(K)$, alors

$$b_A : K^n \times K^n : (u, v) \mapsto uAv^T$$

est une forme bilinéaire sur K^n, appelée la *forme bilinéaire associée à A*.

Démonstration. Pour tous $u, v, w \in E$ et $\alpha, \beta \in K$, on a

$$\begin{array}{rcl} b_A(\alpha u + \beta v, w) & = & (\alpha u + \beta v)Aw^T = (\alpha u)Aw^T + (\beta v)Aw^T \\ & = & \alpha(uAw^T) + \beta v(Aw^t) = \alpha b_A(u, w) + \beta b_A(v, w). \end{array}$$

De même, $b_A(w, \alpha u + \beta v) = \alpha b_A(w, u) + \beta b_A(w, v)$. Ainsi b_A est bilinéaire. Ceci achève la démonstration.

Remarque. Si $A = (a_{ij})_{n\times n}$ alors, pour tous $u = (x_1, \ldots, x_n)$ et $v = (y_1, \ldots, y_n)$, on a

$$uAv^T = \sum\nolimits_{1\le i,j\le n} a_{ij}x_iy_j.$$

Dés maintenant, on étudiera formes bilinéaires sur un espace vectoriel de dimension finie.

9.1.3. Définition. Soit f une forme bilinéaire sur E. Si $\{u_1, \ldots, u_n\}$ est une base de E, la matrice carrée d'ordre n sur K dont le (i,j)-terme est $f(u_i, u_j)$, pour tous $1 \le i, j \le n$, est appelée *matrice de f dans* $\{u_1, \ldots, u_n\}$ et notée $[f]_{\{u_1,\ldots,u_n\}}$.

Exemple. (1) Considérons la forme bilinéaire f sur $\mathbb{R}^2$ donnée par

$$f((x_1, x_2), (y_1, y_2)) = (x_1, x_2) \begin{pmatrix} 1 & -3 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1y_1 - 3x_1y_2 + 4x_2y_1 + 2x_2y_2.$$

Soit $\{e_1, e_2\}$ la base canonique de $\mathbb{R}^2$. Alors

$$f(e_1, e_1) = 1, \ f(e_1, e_2) = -3, \ f(e_2, e_1) = 4, \ f(e_2, e_2) = 2.$$

D'où

$$[f]_{\{e_1,e_2\}} = \begin{pmatrix} 1 & -3 \\ 4 & 2 \end{pmatrix}.$$

(2) Soit b_A la forme bilinéaire associée à $A = (a_{ij}) \in M_n(K)$. Alors A est la matrice de b_A dans la base canonique $\{e_1, \ldots, e_n\}$ de K^n. En effet, pour tous $1 \le i, j \le n$, on a

$$b_A(e_i, e_j) = e_i A e_j^T = a_{ij}.$$

9.1.4. Proposition. Soit $\{u_1, \ldots, u_n\}$ une base de E, et soit f une forme bilinéaire sur E avec $[f]_{\{u_1,\ldots,u_n\}} = (a_{ij})_{n\times n}$.

(1) Si $u = x_1u_1 + \cdots + x_nu_n, \ v = y_1u_1 + \cdots + y_nu_n \in E$, alors

$$f(u, v) = (x_1 \cdots x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

(2) f est symétrique si, et seulement si, A est symétrique.

Démonstration. (1)

$$\begin{aligned} f(u,v) &= f(\textstyle\sum_{i=1}^n x_iu_i, \sum_{j=1}^n y_ju_j) = \sum_{i=1}^n \sum_{j=1}^n x_iy_jf(u_i, u_j) \\ &= \textstyle\sum_{i=1}^n \sum_{j=1}^n x_ia_{ij}y_j = (x_1 \cdots x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \end{aligned}$$

(2) Si f est symétrique, alors $a_{ij} = f(u_i, u_j) = f(u_j, u_i) = a_{ji}$. Ainsi A est symétrique. Réciproquement supposons que $A^T = A$. Alors pour tous $u = x_1u_1 + \cdots + x_nu_n$ et $v = y_1u_1 + \cdots + y_nu_n \in E$,

$$f(u,v) = (x_1 \cdots x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (y_1 \cdots y_n) A^T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (y_1 \cdots y_n) A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = f(v, u).$$

D'où f est symétrique. Ceci s'achève la démonstration.

9.1.5. Proposition. Soit $\{u_1, \ldots, u_n\}$ une base de E. Une fonction $f : E \times E \to K$ est une forme bilinéaire sur E si, et seulement si, il existe une matrice $A = (a_{ij})_{n\times n}$ sur K telle que pour tous $u = x_1u_1 + \cdots + x_nu_n,\ v = y_1u_1 + \cdots + y_nu_n \in E$,

$$f(u,v) = \sum\nolimits_{1\le i,j\le n} a_{ij}x_iy_j = (x_1, \ldots, x_n)A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Dans ce cas, $[f]_{\{u_1,\ldots,u_n\}} = A$.

Démonstration. Il suffit de montrer la suffisance. Posons $A = (a_{ij})_{n\times n}$. Soient $u = \sum_{i=1}^n x_iu_i$, $v = \sum_{i=1}^n y_iu_i$ et $w = \sum_{i=1}^n z_iu_i \in E$ et $\alpha, \beta \in K$. Alors $f(u,v) = (x_1, \ldots, x_n)A(y_1, \ldots, y_n)^T$. Or $\alpha u + \beta v = \sum_{i=1}^n (\alpha x_i + \beta y_i)u_i$. Ainsi

$$\begin{array}{rcl} f(\alpha u + \beta v, w) & = & (\alpha x_1 + \beta y_1, \ldots, \alpha x_n + \beta y_n)A(z_1, \ldots, z_n)^T \\ & = & [\alpha(x_1, \ldots, x_n) + \beta(y_1, \ldots, y_n)]A(z_1, \ldots, z_n)^T \\ & = & \alpha(x_1, \ldots, x_n)A(z_1, \ldots, z_n)^T + \beta(y_1, \ldots, y_n)A(z_1, \ldots, z_n)^T \\ & = & \alpha f(u,w) + \beta f(v,w). \end{array}$$

De même $f(u, \alpha v + \beta w) = \alpha f(u,v) + \beta f(u,w)$. Enfin pour tous $1 \le i,j \le n$, $f(u_i,u_j) = e_iAe_j^T = a_{ij}$. Par conséquent, $[f]_{\{u_1,\ldots,u_n\}} = (a_{ij})_{n\times n}$. Ceci achève la démonstration.

Le résultat suivant nous dit comment sont reliées deux matrices qui représentent la même forme bilinéaire dans deux bases différentes.

9.1.6. Théorème. Soit f une forme bilinéaire sur E. Si $\{u_1, \ldots, u_n\}$ et $\{v_1, \ldots, v_n\}$ sont deux bases de E, alors

$$[f]_{\{v_1,\ldots,v_n\}} = P^T[f]_{\{u_1,\ldots,u_n\}}P,$$

où P est la matrice de passage de $\{u_1, \ldots, u_n\}$ vers $\{v_1, \ldots, v_n\}$.

Démonstration. Posons $P = (p_{ij})_{n\times n} = (P_1P_2\cdots P_n)$ avec P_j la j-ième colonne de P. Alors pour tout $1 \le i,j \le n$, $v_i = p_{1i}u_1 + \cdots + p_{ni}u_n$ et $v_j = p_{1j}u_1 + \cdots + p_{nj}u_n$. Donc

$$f(v_i, v_j) = (p_{1i}\cdots p_{ni})A \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix} = P_i^TAP_j.$$

D'autre part,

$$P^TAP = \begin{pmatrix} P_1^T \\ \vdots \\ P_n^T \end{pmatrix} (AP_1, \cdots, AP_n).$$

Ainsi le (i,j)-terme de P^TAP est également $P_i^TAP_j$. D'où

$$[f]_{\{v_1,\ldots,v_n\}} = (f(v_i,v_j))_{n\times n} = P^TAP.$$

Ceci achève la démonstration.

On dit que deux matrices $A, B \in M_n(K)$ sont *congrues* s'il existe une matrice inversible $P \in M_n(K)$ telle que $B = P^TAP$. Le théorème précédent nous dit que les matrices d'une forme bilinéaire dans deux bases sont congrues.

9.1.7. Proposition. Soit f une forme bilinéaire sur E. Si A est la matrice de f dans une base $\{u_1, \ldots, u_n\}$ de E, alors une matrice $B \in M_n(K)$ est congrue à A si et seulement si, il existe une base $\{v_1, \ldots, v_n\}$ de E telle que $[f]_{\{v_1,\ldots,v_n\}} = B$.

Démonstration. Il suffit de montre la nécesité. Supposons que $B = P^T AP$ avec P inversible. Posons $(v_1, \ldots, v_n) = (u_1, \ldots, u_n)P$. Alors $\{v_1, \ldots, v_n\}$ est une base de E et P est la matrice de passage de $\{u_1, \ldots, u_n\}$ vers $\{v_1, \ldots, v_n\}$. Donc la matrice de f dans $\{v_1, \ldots, v_n\}$ est $P^T AP = B$. Ceci achève la démonstration.

9.2. Formes quadratiques

Dans cette section, nous nous intéressons aux formes bilinéaires symétriques sur les espaces vectoriels réels. On se fixe E un espace vectoriel réel.

9.2.1. Définition. Une fonction $q : E \to \mathbb{R}$ s'appelle une *forme quadratique* sur E s'il existe une forme bilinéaire symétrique f sur E telle que $q(u) = f(u, u)$, pour tout $u \in E$.

Remarque. $q(0_E) = 0$.

Exemple. Si $A = (a_{ij}) \in M_n(\mathbb{R})$ est symétrique, alors A définit une forme bilinéaire symétrique

$$b_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} : (u, v) \mapsto uAv^T.$$

Ainsi

$$q_A : \mathbb{R}^n \mapsto \mathbb{R} : u \mapsto b_A(u, u) = uAu^T$$

est une forme quadratique sur $\mathbb{R}^n$, appelée la *forme quadratique associée à* A.

9.2.2. Définition. Soit q une forme quadratique sur E définie par une forme bilinéaire symétrique f. Si $\{u_1, \ldots, u_n\}$ est une base de E, alors on appelle *matrice de* q *dans* $\{u_1, \ldots, u_n\}$ la matrice de f dans $\{u_1, \ldots, u_n\}$ et on la note $[q]_{\{u_1,\ldots,u_n\}}$.

Remarque. La matrice d'une forme quadratique dans une base est toujours symétrique.

Exemple. Soit $A \in M_n(\mathbb{R})$ symétrique. La matrice de q_A dans la base canonique de $\mathbb{R}^n$ est A.

9.2.3. Proposition. Soit $\{u_1, \ldots, u_n\}$ une base de E. Une fonction $q : E \to \mathbb{R}$ est une forme quadratique sur E si, et seulement si, il existe des nombres réels a_{ij}, $1 \le i \le j \le n$, tels que pour tout $u = x_1u_1 + \cdots + x_nu_n \in E$,

$$q(u) = \sum\nolimits_{i=1}^{n} a_{ii}x_i^2 + \sum\nolimits_{1\le i<j\le n} a_{ij}x_ix_j.$$

Dans ce cas, la matrice de q dans la base $\{u_1, \ldots, u_n\}$ est $B = (b_{ij})_{n\times n}$, où $b_{ii} = a_{ii}$, pour tout $1 \le i \le n$; et $b_{ij} = b_{ji} = \frac{1}{2}\, a_{ij}$, pour tous $1 \le i < j \le n$. En outre, on dit que q est de la *forme diagonale* dans $\{u_1, \ldots, u_n\}$ si $a_{ij} = 0$, pour tous $1 \le i < j \le n$.

Démonstration. Il suffit de montrer la suffisance. D'après le définition, la matrice B est symétrique. Il suit de la proposition 9.1.5 qu'il existe une forme bilinéaire symétrique sur E dont la matrice dans $\{u_1, \ldots, u_n\}$ est B. Or

$$f(u, u) = \sum_{1\le i,j\le n} b_{ij}x_ix_j = \sum_{i=1}^{n} b_{ii}x_i^2 + \sum_{1\le i<j\le n} 2b_{ij}x_ix_j = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{1\le i<j\le n} a_{ij}x_ix_j = q(u).$$

Ceci achève la démonstration.

Le résultat suivant est une conséquence immédiate de 9.1.4 et le théorème 9.1.6.

9.2.4. Théorème. Soit $q : E \to \mathbb{R}$ une forme quadratique sur E et soit A la matrice de q dans une base $\{u_1, \ldots, u_n\}$ de E.

(1) Pour tout $u = x_1u_1 + \cdots + x_nu_n, x_i \in \mathbb{R}$, on a

$$q(u) = (x_1, \ldots, x_n)A(x_1, \ldots, x_n)^T = q_A(x_1, x_2, \ldots, x_n).$$

(2) Si $\{v_1, \ldots, v_n\}$ est une autre base vers laquelle la matrice de passage de $\{u_1, \ldots, u_n\}$ est P, alors

$$[q]_{\{v_1,\ldots,v_n\}} = P^TAP.$$

9.2.5. Définition. Soient f une forme bilinéaire symétrique sur E et q la forme quadratique définie par f.

(1) On dit que q, ainsi que f, est *défine positive* si $q(u) = f(u, u) > 0$, pour tout vecteur non nul u de E.

(2) Une matrice symétrique $A \in M_n(\mathbb{R})$ est dite *défine positive* si q_A est défine positive, c'est-à-dire, $uAu^T > 0$ pour tout vecteur non nul u de $\mathbb{R}^n$.

9.2.6. Proposition. Une forme quadratique (respectivement, une forme bilinéaire symétrique) sur E est définie positive si et seulement si sa matrice dans une base de E est définie positive.

Démonstration. Soient q une forme quadratique et A la matrice de q dans une base $\{u_1, \ldots, u_n\}$. Pour tout $u = x_1u_1 + \cdots + x_nu_n \in E$, $q(u) = q_A(x_1, \ldots, x_n)$. Or u est non nul si et seulement si $(x_1, \ldots, x_n)$ est non nul. Ainsi $q(u) > 0$ pour tout $u \in E$ non nul si et seulement si $q_A(x_1, \ldots, x_n) > 0$ pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n$ non nul. D'où le résultat. Ceci achève la démonstration.

Le résultat précédent nous dit qu'il suffit de discutera la définie positivité de matrices symétriques réelles. On est en train de donner un critère pour qu'une matrice symétrique réelle soit définie positive.

9.2.7. Lemme. Si $A \in M_n(\mathbb{R})$ est symétrique, alors $\chi_A(\lambda)$ est scindé sur $\mathbb{R}$.

Démonstration. Comme $\mathbb{C}$ est algébriquement clos, $\chi_A(\lambda)$ est scindé sur $\mathbb{C}$. Soit $\lambda_0 \in \mathbb{C}$ une valeur propre de A. Alors il existe $u = (x_1, \ldots, x_n) \in \mathbb{C}^n$ non nul tel que $Au^T = \lambda_0u^T$. Ainsi $\lambda_0u = uA$. Posons $\bar{u} = (\bar{x}_1, \ldots, \bar{x}_n)$, où $\bar{x}_i$ est le conjugué de x_i. Alors $A\bar{u}^T = \overline{Au^T} = \overline{\lambda_0u^T} = \bar{\lambda}_0\bar{u}^T$. De plus, $u\bar{u}^T = \sum_{i=1}^n x_i\bar{x}_i = \sum_{i=1}^n ||x_i||^2 > 0$ car $u \neq 0$. Or

$$\lambda_0(u\bar{u}^T) = (uA)\bar{u}^T = u(\bar{\lambda}_0\bar{u}^T) = \bar{\lambda}_0(u\bar{u}^T).$$

Par conséquent, $\bar{\lambda}_0 = \lambda_0$, c'est-à-dire, λ_0 est réel. Ceci achève la démonstration.

9.2.8. Corollaire. Si $A \in M_n(\mathbb{R})$ est définie positive, alors les valeurs propres de A dans $\mathbb{C}$ sont toutes réelles positives et $\det(A) > 0$.

Démonstration. Soit λ_0 une valeur propre de A. Alors λ_0 est réelle d'après le lemme 9.2.7. Donc il existe $u \in \mathbb{R}^n$ non nul tel que $Au^T = \lambda_0u^T$. Comme A est définie positive, $uAu^T = \lambda_0uu^T$ est positif. Ainsi $\lambda_0 > 0$ car $uu^T > 0$. Enfin $\det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$ avec les λ_i sont réels positifs. Ainsi $\det(A) = \lambda_1 \cdots \lambda_n > 0$. Ceci achève la démonstration.

9.2.9. Lemme. Soient $A, B \in M_n(\mathbb{R})$ symétriques. Si A, B sont congrues, alors A est définie positive si et seulement si B l'est.

Démonstration. Par h'hypothèse, $B = P^TAP$ avec P inversible. Supposons que A est définie positive. Comme $B^T = P^TA^TA = P^TAP = B$, on voit que B est symétrique. Si $u \in \mathbb{R}^n$ est non nul, alors $uP^T \in \mathbb{R}^n$ est non nul. Ainsi $(uP^T)A(Pu^T) > 0$, c'est-à-dire, $uBu^T > 0$. Donc B est définie positive.

Si B est définie positive, comme $A = (P^{-1})^TAP^{-1}$, alors A est définie positive. Ceci achève la démonstration.

Soit $A = (a_{ij})_{n\times n}$ une matrice carrée sur un corps. Pour tout $1 \le r \le n$, le déterminant de $(a_{ij})_{1\le i\le r;1\le j\le r}$ s'appelle *mineur principal* d'ordre r de A. En particulier, le mineur principal d'ordre n est $\det(A)$.

9.2.10. Théorème. Une matrice réelle symétrique est définie positive si, et seulement si, tous ses mineurs principaux sont positifs.

Démonstration. Posons $A = (a_{ij})_{n\times n}$. Procédons par récurrence en n. Si $n = 1$, il est évident que A est définie positive si et seulement si $a_{11} > 0$. Supposons que $n > 1$ et le résultat est vrai pour $n - 1$. Partageons A comme suit:

$$A = \begin{pmatrix} B & C \\ C^T & a_{nn} \end{pmatrix}, \text{ où } B = (a_{ij})_{1\le i\le n-1;1\le j\le n-1},\ C = (a_{in})_{1\le i\le n-1}.$$

Supposons premièrement que A est définie positive. D'après le corollaire 9.2.8, $\det(A) > 0$. On prétend que B est également définie positive. En effet, B est évidemment symétrique. Pour tout $x \in \mathbb{R}^{n-1}$ non nul, on a

$$xBx^T = (x\ \ 0)\begin{pmatrix} B & C \\ C^T & a_{nn} \end{pmatrix}\begin{pmatrix} x^T \\ 0 \end{pmatrix} = (x\ \ 0)A\begin{pmatrix} x^T \\ 0 \end{pmatrix} > 0.$$

Donc B est définie positive. Par hypothèse de récurrence, tous les mineurs principaux de B sont positifs. Ainsi tous mineurs principaux de A sont positifs.

Supposons réciproquement que tous les mineurs principaux de A sont positifs. C'est-à-dire, $\det(A) > 0$ et tous les mineurs principaux de B sont positifs. Par hypothèse de récurrence, B est définie positive. Remarquons que B est inversible car $\det(B) > 0$. Posons

$$Q_{n\times n} = \begin{pmatrix} I_{n-1} & -B^{-1}C \\ 0 & 1 \end{pmatrix}.$$

Alors $\det(Q) = 1$ et

$$Q^TAQ = \begin{pmatrix} I_{n-1} & 0 \\ -C^TB^{-1} & 1 \end{pmatrix}\begin{pmatrix} B & C \\ C^T & a_{nn} \end{pmatrix}\begin{pmatrix} I_{n-1} & -B^{-1}C \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} B & 0 \\ 0 & b \end{pmatrix}, \quad b \in \mathbb{R}.$$

Comme $b\det(B) = \det(Q^TAQ) = \det(Q)\det(A)\det(Q) = \det(A) > 0$ et $\det(B) > 0$, on a $b > 0$. Si $u \in \mathbb{R}^n$ est non nul, alors $u = (x\ x_n)$ avec $x \in \mathbb{R}^{n-1}$ et $x_n \in \mathbb{R}$. Or $u(Q^TAQ)u^T = xBx^T + bx_n^2$. Si x est non nul, alors $xBx^T > 0$. Si x est nul, alors $x_n \ne 0$, et donc $bx_n^2 > 0$. Ainsi $u(Q^TAQ)u^T > 0$ en tout cas. Ceci montre que Q^TAQ est définie positive. D'après le lemme 9.2.9, A est définie positive. Ceci achève la démonstration.

Exemple. Une matrice diagonale réelle est définie positive si, et seulement si, les termes diagonaux sont tous positifs.

9.3. Produit scalaire

Partout dans cette section, on se fixe E un espace réel non nul. Un produit scalaire sur E nous permet d'introduire les longueurs, les distances et les angles de vecteurs de E, c'est ce que l'on appelle la *structure métrique* de E.

9.3.1. Définition. Une forme bilinéaire sur E

$$< \, , \, >: E \times E \mapsto \mathbb{R} : (u, v) \mapsto < u, v >$$

s'appelle *produit scalaire*si elle est symétrique et définie positive.

Si E est de dimension finie muni d'un produit scalaire, on dit alors que E est un *espace euclidien* par rapport à ce produit scalaire.

Exemple. (1) L'espace réel $\mathbb{R}^n$ est un espace euclidien pour le produit scalaire canonique, ce qui est défini comme suit:

$$< (x_1, \ldots, x_n), \, (y_1, \ldots, y_n) >= x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

En effet, la matrice de $< \, , \, >$ dans la base canonique est I_n, ce qui est symétrique et définie positive.

(2) De même, l'espace réel $\mathbb{R}^{(n)}$ est un espace euclidien pour le produit scalaire canonique, ce qui est défini comme suit:

$$< (x_1, \ldots, x_n)^T, \, (y_1, \ldots, y_n)^T >= x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

(3) L'espace réel $\mathbb{R}_n[x]$ est un espace euclidien pour le produit scalaire canonique, ce qui est défini comme suit:

$$< a_1 + a_2x + \cdots + a_nx^{n-1}, \, b_1 + b_2x + \cdots + b_nx^{n-1} >= a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

En effet, il s'agit d'une forme binileáire dont la matrice dans la base canonique est I_n.

(4) Soient $a, b \in \mathbb{R}$ avec $a < b$. Considérons l'espace réel $C[a, b]$ des fonctions continues $f : [a, b] \mapsto \mathbb{R}$. Alors la forme bilinéaire $< \, , \, >$ définie par

$$< f \, , \, g >= \int_a^b f(t)g(t)dt$$

est un produit scalaire, appelé *produit scalaire canonique* sur $C[a, b]$.

En effet, on a vu que $< \, , \, >$ est bilinéaire et symétrique. Soit $f \in C[a, b]$ non nul. Alors il existe un $t_0 \in [a, b]$ tel que $f(t_0) \neq 0$, et donc $f^2(t_0) > 0$. Comme f^2 est continue, il existe un interval $[c, d] \subseteq [a, b]$ avec $c < d$ et un constant $q > 0$ tels que $f^2(t) \geq q$, pour tout $t \in [c, d]$. Or

$$< f, f >= \int_a^b f^2(t)dt \geq \int_c^d f^2(t)dt \geq \int_c^d qdt = q(d - c) > 0.$$

Ainsi $< \, , \, >$ est définie positive, et donc un produit scalaire sur $C[a, b]$.

Dorénavant, lorsque nous parlerons les espaces euclidiens $\mathbb{R}^n$, $\mathbb{R}^{(n)}$, $\mathbb{R}_n[x]$ ou $C[a, b]$ sans autres spécifications, alors le produit scalaire sera le produit scalaire canonique.

Dès maintenant, on se fixe $< \, , \, >$ un produit scalaire sur E.

9.3.2. Définition. Si $u \in E$, le nombre réel non négatif

$$||u|| = \sqrt{< u, u >}$$

s'appelle la *longueur* ou la *norme* de u.

Remarque. (1) Pour tout $\alpha \in \mathbb{R}$, $||\alpha u|| = |\alpha| \cdot ||u||$.
(2) Comme le produit scalaire est défini positif, $||u|| = 0$ si et seulement si $u = 0_E$.

Exemple. (1) Dans $\mathbb{R}^n$, on a $||(x_1, x_2, \ldots, x_n)|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.
(2) Pour toute $f \in C[a, b]$, on a

$$||f|| = \sqrt{\int_a^b f^2 t dt}.$$

9.3.3. Inégalité de Cauchy-Schwarz. Si $u, v \in E$, alors

$$| < u, v > | \leq ||u|| \cdot ||v||.$$

En outre, l'égalité a lieu si et seulement si u et v sont linéairement dépendants.

Démonstration. Soient $u, v \in E$. Si $v = 0_E$, le résultat est trivial. Supposons que $v \neq 0_E$. Alors $< v,\ v >> 0$. Or

$$\begin{aligned} 0 \quad &\leq \quad < u - \frac{< u, v >}{< v, v >} v, u - \frac{< u, v >}{< v, v >} v > = < u, u > - \frac{< u, v >^2}{< v, v >} \\ &= \quad \frac{< u, u >< v, v > - < u, v >^2}{< v, v >}. \end{aligned}$$

D'où $< u, v >^2 \leq < u, u >< v, v >= ||u||^2\, ||v||^2$, ce qui donne $| < u, v > | \leq ||u|| \cdot ||v||$.

Si $| < u, v > | = ||u|| \cdot ||v||$, alors

$$< u - \frac{< u, v >}{< v, v >} v, u - \frac{< u, v >}{< v, v >} v >= 0. \quad \text{Donc} \quad u - \frac{< u, v >}{< v, v >} v = 0_E,$$

d'où $\{u, v\}$ est liée. Réciproquement si $\{u, v\}$ est liée, alors $u = \alpha v$ avec $\alpha \in \mathbb{R}$. Or

$$\begin{aligned} < u, v >^2 \quad &= \quad < \alpha v, v >< \alpha v, v >=< \alpha v, v > \alpha < v, v > \\ &= \quad < \alpha v, \alpha v >< v, v >=< u, u >< v, v > . \end{aligned}$$

Par conséquent, $| < u, v > | = ||u|| \cdot ||v||$. Ceci achève la démonstration.

9.3.4. Corollaire. (1) Pour tous $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$,

$$(x_1 y_1 + \cdots + x_n y_n)^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2).$$

(2) Pour tous $f, g \in C[a, b]$,

$$\left(\int_a^b f(t) g(t) dt \right)^2 \leq \int_a^b f^2(t) dt \int_0^1 g^2(t) dt.$$

9.3.5. Inégalité triangulaire. Si $u, v \in E$, alors $||u + v|| \leq ||u|| + ||v||$.
Démonstration.

$$\begin{aligned} &||u + v||^2 =< u + v, u + v >=< u, u > + < u, v > + < v, u > + < v, v > \\ = \quad &||u||^2 + 2 < u, v > + ||v||^2 \leq ||u||^2 + 2||u|| \cdot ||v|| + ||v||^2 = (||u|| + ||v||)^2. \end{aligned}$$

Donc $||u+v|| \leq ||u|| + ||v||$. Ceci achève la démonstration.

9.3.6. Définition. Soient u et v des vecteurs de E. On appelle *distance* entre u et v, notée $d(u,v)$, la longueur de $u-v$.

Remarque. Soient u, v, et $w \in E$.
(1) $d(u,v) = 0$ si et seulement si $u = v$.
(2) $d(u,w) \leq d(u,v) + d(v,w)$.

9.3.7. Définition. Soient $u, v \in E$ tous non nuls. D'après l'inégalité de Cauchy-Schwarz,

$$\frac{|<u,v>|}{||u|| \cdot ||v||} \leq 1.$$

Par conséquent, il existe un unique $\theta \in [0, \pi]$ tel que

$$\cos\theta = \frac{<u,v>}{||u|| \cdot ||v||}.$$

On dit que θ est l'*angle* entre u et v. On a ainsi

$$<u,v> = ||u|| \cdot ||v|| \cos\theta.$$

Exemple. Soit $\{e_1, \ldots, e_n\}$ la base canonique de $\mathbb{R}^n$. Si $i \neq j$, alors l'angle entre e_i et e_j est $\frac{\pi}{2}$.

On verra que la notion d'angle définie ici est une généralisation de la notion d'angle usuelle dans le plan. En effet, soient $u_1 = (x_1, y_1),\ u_2 = (x_2, y_2) \in \mathbb{R}^2$ tous non nuls. Si ϕ_i est l'angle entre l'axe des x et le vecteur u_i, alors

$$\cos\phi_i = \frac{x_i}{\sqrt{x_i^2 + y_i^2}}, \quad \sin\phi_i = \frac{y_i}{\sqrt{x_i^2 + y_i^2}}, \quad i = 1, 2.$$

Supposons que $\phi_1 > \phi_2$. Alors l'angle θ entre u_1 et u_2 est égal à $\phi_1 - \phi_2$. Par conséquent,

$$\begin{aligned}\cos\theta &= \cos(\phi_1 - \phi_2) = \cos\phi_1 \cos\phi_2 + \sin\phi_1 \sin\phi_2 \\ &= \frac{x_1}{\sqrt{x_1^2 + y_1^2}} \frac{x_2}{\sqrt{x_2^2 + y_2^2}} + \frac{y_1}{\sqrt{x_1^2 + y_1^2}} \frac{y_2}{\sqrt{x_2^2 + y_2^2}} = \frac{<u_1, u_2>}{||u_1||\, ||u_2||}.\end{aligned}$$

9.4. Orthogonalité

Partout dans cette section, on se fixe E un espace réel muni d'un produit scalaire $<\ ,\ >$.

9.4.1. Définition. Deux vecteurs u et v de E sont dits *orthogonaux* si $<u,v> = 0$.

Remarque. (1) Pour tout $u \in E$, u et 0_E sont toujours orthogonaux.
(2) Si u et v sont tous non nuls, alors u et v sont orthogonaux si et seulement si l'angle entre u et v est $\dfrac{\pi}{2}$.

Le résultat suivant est trivial.

9.4.2. Proposition. Soit F un sous-espace de E.
(1) $F^{\perp} = \{u \in E \mid \ < u, v >= 0, \ \text{ pour tout } \ v \in F\}$ est un sous-espace de E, appelé *orthogonal* de F.
(2) Si F est engendré par $u_1, \ldots, u_r$, alors $u \in F^{\perp}$ si et seulement si $< u, \ u_i >= 0$, $i = 1, \ldots, r$.

9.4.3. Définition. Une famille $\{u_1, \ldots, u_r\}$ de vecteurs de E est dite *orthogonale* si $< u_i, u_j >= 0$ pour tous $1 \leq i, j \leq r$ avec $i \neq j$; et *orthonormée* si $< u_i, u_j >= \delta_{ij}$, pour tous $1 \leq i, j \leq r$.

Remarque. (1) La famille $\{u_1, \ldots, u_r\}$ est orthonormée si elle est orthogonale et chaque vecteur est de longueur 1.
(2) Une *base orthogonale* (respectivement, *orthonormée*) de E est une base qui est une famille orthogonale (respectivement, orthonormée).
(3) Une famille d'un seul vecteur est toujours orthogonale.

On voit aisément que les bases canoniques de $\mathbb{R}^n$, $\mathbb{R}^{(n)}$ et $\mathbb{R}_n[x]$ sont des bases orthonormées par rapport au produit scalaire canonique. Réciproquement, on a le résultat suivant.

9.4.4. Proposition. Soit E un espace euclidien muni d'un produit scalaire $<\, , >$. Une base $\{u_1, \ldots, u_n\}$ de E est orthonormée si, et seulement si, la matrice de $<\, , >$ dans $\{u_1, \ldots, u_n\}$ est I_n. Dans ce cas,
(1) Pour tous $u = x_1u_1 + \cdots + x_nu_n$ et $v = y_1u_1 + \cdots y_nu_n$, on a

$$< u, v >= x_1y_1 + \cdots + x_ny_n.$$

(2) Pour tout $u \in E$, on a $u =< u, u_1 > u_1 + \cdots + < u, u_n > u_n$.
Démonstration. La partie (1) est triviale. Pour montrer (2), supposons que $u = \sum_{i=1}^n \alpha_i u_i, \ \alpha_i \in \mathbb{R}$. Alors pour tout $1 \leq j \leq n$,

$$< u, u_j >= \sum_{i=1}^n \alpha_i < u_i, \ u_j >= \sum_{i=1}^n \alpha_i \delta_{ij} = \alpha_j.$$

La preuve s'achève.

Exemple. Soit $\{e_1, \ldots, e_n\}$ la base canonique de $\mathbb{R}^n$. Pour $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $x_i =< e_i, x >$, $i = 1, \ldots, n$.

On est en train de montrer que tout espace euclidien admet une base orthonormée.

9.4.5. Lemme. Soit $\{u_1, \ldots, u_r\}$ une famille orthogonale de vecteurs non nuls de E.
(1) La famille $\{u_1, \ldots, u_r\}$ est libre.
(2) La famille $\{\alpha_1 u_1, \ldots, \alpha_r u_r\}$ avec $\alpha_i \in \mathbb{R}$ est également orthogonale.
(3) La famille $\left\{\dfrac{u_1}{||u_1||}, \ldots, \dfrac{u_r}{||u_r||}\right\}$ est orthonormée.

Démonstration. (1) Supposons que $\beta_1 u_1 + \cdots + \beta_r u_r = 0_E, \ \alpha_i \in \mathbb{R}$. Alors pour tout $1 \leq i \leq r$, on a $0 =< u_i, 0_E >=< u_i, \sum_{j=1}^n \beta_j u_j >= \sum_{j=1}^n \beta_j < u_i, u_j >= \beta_i < u_i, u_i >$. Ceci donne $\beta_i = 0$ car $< u_i, u_i > \neq 0$. Ainsi $\{u_1, \ldots, u_r\}$ est libre.
(2) Si $i \neq j$, alors $< \alpha_i u_i, \alpha_j u_j >= \alpha_i \alpha_j < u_i, u_j >= 0$.
(3) D'après la partie (2), la famille

$$\left\{\frac{u_1}{||u_1||}, \ldots, \frac{u_r}{||u_r||}\right\}$$

est orthogonale. Or pour tout $1 \le i \le r$,

$$< \frac{u_i}{||u_i||}, \frac{u_i}{||u_i||} >= \frac{1}{||u_i|| \cdot ||u_i||} < u_i, u_i >= \frac{||u_i||^2}{||u_i||^2} = 1.$$

La preuve s'achève.

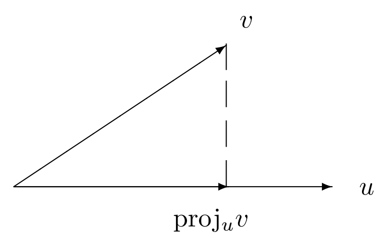
Exemple. On a vu que $\{1, \cos t, \sin t\}$ est orthogonale dans $C[0, 2\pi]$. Donc $\{1, \cos t, \sin t\}$ est libre dans $C[0, 2\pi]$.

9.4.6. Corollaire. Si $\dim(E) = n$, alors une famille orthogonale de n vecteurs non nuls de E est une base orthogonale de E.

9.4.7. Définition. Soit $u \in E$ non nul. Pour tout $v \in E$, on appelle *projection* de v sur u le vecteur

$$\text{proj}_u v = \frac{< u, v >}{< u, u >} u,$$

ce qui est illustré par le diagramme suivant:



Remarque. On voit aisément que $v - \text{proj}_u v$ est orthogonal à u.

Exemple. Soit $\{e_1, \ldots, e_n\}$ la base canonique de $\mathbb{R}^n$. Pour tous $u = (x_1, \ldots, x_n)$, $\text{proj}_{e_i} u = x_i e_i$, $i = 1, \ldots, n$.

9.4.8. Définition. Soient F un sous-espace vectoriel non nul de E admettant une base orthogonale $\{u_1, \ldots, u_r\}$. Pour tout $v \in E$, on appelle *projection orthogonale de v sur F* le vecteur tel que défini ci-dessous:

$$\text{proj}_F v = \text{proj}_{u_1} v + \cdots + \text{proj}_{u_r} v \in F.$$

9.4.9. Proposition. Soit F un sous-espace vectoriel de E. Pour tout $v \in E$, on a
(1) $v - \text{proj}_F v$ appartient à $F^\perp$.
(2) $d(v, \text{proj}_F v) < d(v, u)$, pour tout $u \in F$ différent de $\text{proj}_F v$.
Démonstration. (1) Pour tout $1 \le i \le r$, on a

$$< v - \text{proj}_F v,\ u_i >=< v, u_i > - \sum_{j=1}^{r} \frac{< v, u_j >}{< u_j, u_j >} < u_j, u_i >=< v, u_i > - < v, u_i >= 0.$$

Par conséquent, $v - \text{proj}_F v \in F^\perp$.

(2) Soit $u \in F$. Écrivons $v - u = (v - \text{proj}_F v) + (\text{proj}_F v - u)$. Comme $u, \text{proj}_F v \in F$, on voit que $v - \text{proj}_F v$ et $\text{proj}_F v - u$ sont orthogonaux. Ainsi

$$||v - u||^2 = ||v - \text{proj}_F v||^2 + ||\text{proj}_F v - u||^2.$$

Si $u \neq \text{proj}_F v$, alors $||\text{proj}_F v - u|| > 0$. D'où $||v - u||^2 > ||v - \text{proj}_F v||^2$. Par conséquent, $d(v, \text{proj}_F v) < d(v, u)$. Ceci achève la démonstration.

Remarque. (1) À cause de la propriété énoncée dans la proposition 9.4.9(2), $\text{proj}_F v$ s'appelle la *meilleure approximation* de v par des vecteurs de F.
(2) $\text{proj}_F v$ est indépendant du choix de base orthogonale de F.

Voici la méthode pour trouver une base orthogonale et une base orthonormée.

9.4.10. Procédé d'orthogonalisation de Gram-Schmidt. Soit E un K-espace vectoriel muni d'un produit scalaire. Soit F un sous-espace vectoriel de E ayant pour base $\{u_1, \ldots, u_r\}$. On pose successivement

$$\begin{array}{rcl} v_1 & = & u_1, \\ v_2 & = & u_2 - \text{proj}_{\,v_1} u_2, \\ v_3 & = & u_3 - \text{proj}_{\,v_1} u_3 - \text{proj}_{\,v_2} u_3, \\ & \vdots & \\ v_r & = & u_r - \text{proj}_{\,v_1} u_r - \text{proj}_{\,v_2} u_r - \cdots - \text{proj}_{\,v_{r-1}} u_r. \end{array}$$

Alors $\{v_1, \ldots, v_r\}$ est une base orthogonale de F. De plus, posant

$$w_i = \frac{v_i}{||v_i||}, \; i = 1, \ldots, r,$$

on obtient une base orthonormée $\{w_1, \ldots, w_r\}$ de F.

Démonstration. Posons $F_p =< u_1, \ldots, u_p >$, et donc $\dim F_p = p$, $p = 1, \cdots, r$. On prétend que $\{v_1, \ldots, v_p\}$ est une base orthogonale de F_p pour tout $1 \leq p \leq r$. En effet, c'est vrai pour $p = 1$. Supposons que l'énoncé est vrai pour p avec $1 \leq p < r$. Or $v_{p+1} = u_{p+1} - \text{proj}_{F_p} u_{p+1} \in F_{p+1}$, et d'après la proposition 9.4.9(1), $v_{p+1} \in F_p^{\perp}$. En particulier, $\{v_1, \ldots, v_p, v_{p+1}\}$ est orthogonale. De plus, $v_{p+1} \neq 0$ car $u_{p+1} \notin F_p$. Ainsi $\{v_1, \ldots, v_p, v_{p+1}\}$ est une base orthogonale de F_{p+1}. Ceci achève la démonstration.

9.4.11. Corollaire. Un espace euclidien admet toujours une base orthonormée.

9.4.12. Théorème. Si F est un sous-espace de E, alors $E = F \oplus F^{\perp}$. En particulier,

$$\dim(E) = \dim(F) + \dim(F^{\perp}).$$

Démonstration. Soit $\{u_1, \ldots, u_r\}$ une base de F, qui se prolonge dans une base $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ de E. Soit $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ la base orthonormée de E obtenue en appliquant le procédé de Gram-Schmidt à $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$. Comme on a vu dans la démonstration du théorème 9.4.10, $\{v_1, \ldots, v_r\}$ est une base de F. Comme $< v_j, v_i >= 0$ pour tout $j > i$, on voit que $\{v_{r+1}, \ldots, v_n\}$ est une famille libre de vecteurs de $F^{\perp}$. En outre, pour tout $v \in F^{\perp}$, d'après la proposition 9.4.4(2),

$$v = \sum\nolimits_{i=1}^{n} < v, v_i > v_i = \sum\nolimits_{i=r+1}^{n} < v, v_i > v_i.$$

D'où $\{v_{r+1}, \ldots, v_n\}$ est une base de $F^{\perp}$. Ceci achève la démonstration.

9.5. Endomorphismes adjoints

Partout dans cette section, on se fixe E un espace euclidien muni d'un produit scalaire $< , >$. Si $u \in E$, on voit aisément que

$$u^* : E \to \mathbb{R} : v \mapsto < u, v >$$

est une forme linéaire sur E.

9.5.1. Lemme. L'application

$$\Phi : E \to E^* : u \mapsto u^*$$

est un isomorphisme de K-espaces vectoriels.

Démonstration. Il est évident que Φ est linéaire. Si $u \in \ker(\Phi)$, alors $< u, v >= 0$ pour tout $v \in E$. En particulier, $< u, u >= 0$. Comme $< , >$ est définie positive, $u = 0_E$. Donc Φ est injective. Comme E et E^* sont de même dimension finie, on a que Φ est un isomorphisme. Ceci achève la démonstration.

9.5.2. Théorème. Pour $T \in$End$\,(E)$, il existe un unique $T^* \in$End$\,(E)$, appelé l'*adjoint* de T, tel que pour tous $u, v \in E$,

$$< T(u),\ v >=< u,\ T^*(v) > .$$

Démonstration. Soit $T \in$ End$\,(E)$. Alors $T^t \in$End$\,(E^*)$ est tel que $T^t(g) = gT$ pour tout $g \in E^*$. Comme $\Phi : E \to E^*$ est un isomorphisme, $T^* = \Phi^{-1}T^t\Phi \in$End$\,(E)$ est tel que $T^t\Phi = \Phi T^*$. Ainsi pour tout $v \in E,\ \Phi(T^*(v)) = T^t(\Phi(v)) = \Phi(v)T$. Donc pour tous $u, v \in E$,

$$< T(u), v >= v^*(T(u)) = \Phi(v)(T(u)) = (\Phi(v)T)(u) = \Phi(T^*(v))(u) =< u, T^*(v) > .$$

Enfin soit $S \in$End$\,(E)$ tel que $< u, S(v) >=< T(u), v >=< u, T^*(v) >$ pour tous $u, v \in E$. Alors pour tout $v \in E$, $S(v) = T^*(v)$. D'où $S = T^*$. Ceci achève la démonstration.

Exemple. On voit que $\mathbf{1}^* = \mathbf{1}$ car $< \mathbf{1}(u), v >=< u, \mathbf{1}(v) >$ pour tous $u, v \in E$.

9.5.3. Théorème. Soit $T \in$ End$\,(E)$ dont T^* est l'adjoint. Si $\{u_1, \ldots, u_n\}$ est une base orthonormée de E, alors $[T^*]_{\{u_1,\ldots,u_n\}}$ est égal à la transposée de $[T]_{\{u_1,\ldots,u_n\}}$.

Démonstration. Posons $[T]_{\{u_1,\ldots,u_n\}} = (a_{ij})_{n\times n}$ et $[T^*]_{\{u_1,\ldots,u_n\}} = (b_{ij})_{n\times n}$. On a alors $T(u_i) = \sum_{j=1}^n a_{ji}u_j,\ i = 1, \ldots, n$; et $T^*(u_j) = \sum_{i=1}^n b_{ij}u_i,\ j = 1, \ldots, n$. D'après la proposition 9.4.4(2), on a

$$b_{ij} =< T^*(u_j), u_i >=< u_j, T(u_i) >=< T(u_i), u_j >= a_{ji}.$$

Ceci achève la démonstration.

9.5.4. Définition. On dit que $T \in$End$\,(E)$ est *auto-adjoint* si $T^* = T$, c'est-à-dire, pour tous $u, v \in E$,

$$< T(u), v >=< u, T(v) > .$$

Exemple. L'endomorphisme identité $\mathbf{1}_E$ et l'endomorphisme nul $\mathbf{0}$ sont tous auto-adjoints.

9.5.5. Proposition. Soit $T \in$ End$\,(E)$. Si $\{u_1, \ldots, u_n\}$ est une base orthonormée de E, alors T est auto-adjoint si et seulement si $[T]_{\{u_1,\ldots,u_n\}}$ est symétrique.

Démonstration. D'aprè le théorème 9.5.3, $[T^*] = [T]^T$. Donc $T^* = T$ si et seulement si $[T^*] = [T]$ si et seulement si $[T]^T = [T]$. Ceci achève la démonstration.

Exemple. Soit $A \in M_n(\mathbb{R})$. Alors

$$T : \mathbb{R}^n \to \mathbb{R}^n : (x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n)A$$

est auto-adjoint si et seulement si A est symétrique.

9.6. Matrices orthogonales et endomorphismes orthogonaux

9.6.1. Définition. Une matrice $A \in M_n(\mathbb{R})$ est dite *orthogonale* si $A^T A = AA^T = I_n$, c'est-à-dire $A^T = A^{-1}$.

Remarque. (1) A est orthogonale si et seulement si A^T l'est.
(2) Si $A^T A = I_n$ ou $AA^T = I_n$, alors $A^{-1} = A^T$. Donc A est orthogonale.
(3) Si A est orthogonale, alors $\det(A) = \pm 1$. Si $\det(A) = 1$, alors A est dite *orthogonale spéciale.*

Exemple. I_n est orthogonale car $I_n^T I_n = I_n I_n = I_n$.

Dès maintenant, on se fixe E un espace euclidien muni d'un produit scalaire $< , >$.

9.6.2. Proposition. Soit $\{u_1, \ldots, u_n\}$ une base orthonormée de E. Alors n vecteurs $v_1, \ldots, v_n$ forment une base orthonormée de E si et seulement si $P^{\{v_1,\ldots,v_n\}}_{\{u_1,\ldots,u_n\}}$ est orthogonale.

Démonstration. On peut supposer que $P = P^{\{v_1,\ldots,v_n\}}_{\{u_1,\ldots,u_n\}}$ est inversible, et donc $\{v_1, \ldots, v_n\}$ est une base de E dans laquelle la matrice du produit scalaire $<,>$ est $P^T I_n P = P^T P$. Ainsi $\{v_1, \ldots, v_n\}$ est une base orthonormée si et seulement si $P^T P = I_n$ si, et seulement si, P est orthogonale. La preuve s'achève.

Remarque. La matrice de passage d'une base orthonormée à une autre base orthonormée est orthogonale.

Remarquons qu'une matrice $A \in M_n(\mathbb{R})$ est la matrice des coordonnées de ses colonnes dans la base canonique de $\mathbb{R}^{(n)}$ et A^T est la matrice des coordonnées de ses lignes dans la base canonique de $\mathbb{R}^n$. En appliquant 9.6.2, on obtient le résultat suivant.

9.6.3. Théorème. Une matrice de $M_n(\mathbb{R})$ est orthogonale si et seulement si ses colonnes forment une base orthonormée de $\mathbb{R}^{(n)}$ si et seulement si ses lignes forment une base orthonormée de $\mathbb{R}^n$.

9.6.4. Définition. Un endomorphisme T de E est dit *orthogonal* si pour tous $u, v \in E$,

$$< T(u), T(v) >=< u, v > .$$

Remarque. Si T est orthogonal, alors $T(u) \neq 0_E$ pour tout $u \neq 0_E$. Donc T est injectif, et ainsi un automorphisme de E car E est de dimension finie.

Exemple. $\mathbb{1} : E \to E : u \mapsto u$ et $-\mathbb{1} : E \to E : u \mapsto -u$ sont orthogonaux.

9.6.5. Lemme. Soit $\{u_1, \ldots, u_n\}$ une base de E. Alors $T \in \text{End}(E)$ est orthogonal si et seulement si $< T(u_i), T(u_j) >=< u_i, u_j >$, pour tous $1 \le i, j \le n$.

Démonstration. Supposons que $< T(u_i), T(u_j) >=< u_i, u_j >$, pour tous $1 \le i, j \le n$. Alors pour tous $u = \sum_{i=1}^n \alpha_i u_i, v = \sum_{i=1}^n \beta_j u_j \in E$,

$$\begin{aligned} < T(u), T(v) > &= < \textstyle\sum_{i=1}^n \alpha_i T(u_i),\ \sum_{i=1}^n \beta_j T(u_j) >= \sum_{1\le i,j\le n}^n\ \alpha_i\beta_j < T(u_i), T(u_j) > \\ &= \textstyle\sum_{1\le i,j\le n}^n\ \alpha_i\beta_j < u_i, u_j >=< u, v > . \end{aligned}$$

Ainsi T est orthogonal.

9.6.6. Théorème. Soit $T \in \text{End}(E)$. Les conditions suivantes sont équivalentes:
(1) T est orthogonal.
(2) $||T(u)|| = ||u||$, pour tout $u \in E$.
(3) $d(u, v) = d(T(u), T(v))$, pour tous $u, v \in E$.

Démonstration. Supposons que T est orthogonal. Pour tout $u \in E$,

$$||u|| = \sqrt{< u, u >} = \sqrt{< T(u), T(u) >} = ||T(u)||.$$

Supposons maintenant que $||T(u)|| = ||u||$, pour tout $u \in E$. Donc pour tous $u, v \in E$,

$$d(T(u), T(v)) = ||T(u) - T(v)|| = ||T(u - v)|| = ||u - v|| = d(u, v).$$

Enfin supposons que $d(T(u), T(v)) = d(u, v)$, c'est-à-dire, $||T(u-v)|| = ||u-v||$, pour tous $u, v \in E$. En particulier, $||T(u)|| = ||u||$ pour tout $u \in E$. Or $< T(u) - T(v),\ T(u) - T(v) >=< u - v,\ u - v >$. Ceci qui donne

$$< T(u), T(u) > -2 < T(u), T(v) > + < T(u), T(v) >=< u, u > -2 < u, v > + < v, v > .$$

Par conséquent, $< T(u), T(v) >=< u, v >$. Ceci montre que T est orthogonal. La preuve s'achève.

9.6.7. Corollaire. Soit $T \in \text{End}(E)$ orthogonal. Si $u, v \in E$ sont tous non nuls, alors l'angle entre u et v est égal à l'angle entre $T(u)$ et $T(v)$.

Démonstration. Soient θ l'angle entre u et v, et ϕ l'angle entre $T(u)$ et $T(v)$. Alors

$$\cos\theta = \frac{< u, v >}{||u|| \cdot ||v||} = \frac{< T(u), T(v) >}{||T(u)|| \cdot ||T(v)||} = \cos\phi.$$

D'où, $\theta = \phi$. Ceci achève la démonstration.

Remarque. La réciproque du corollaire 9.6.7 n'est pas valide.

9.6.8. Théorème. Un endomorphisme T de E est orthogonal si, et seulement si, T est un automorphisme et $T^{-1} = T^*$.

Démonstration. Supposons que T est orthogonal. On se fixe $v \in E$. Alors pour tout $u \in E$, on a $< u, v >=< T(u), T(v) >=< u, (T^*T)(v) >$. Ainsi $(T^*T)(v) = (v)$, pour tout $v \in E$. Donc $T^*T = \mathbf{1}_E$. Comme T est inversible, on a $T^* = T^{-1}$. Supposons maintenant que $T^* = T^{-1}$. Alors $T^*T = \mathbf{1}_E$. Pour tous $u, v \in E$, on a

$$< u, v >=< u,\ T^*(Tv) >=< T(u),\ T(v) > .$$

D'où T est orthogonal. Ceci achève la démonstration.

9.6.9. Proposition. Soit $T \in \text{End}(E)$ et soit $\{u_1, \ldots, u_n\}$ une base orthonormée de E. Les conditions suivantes sont équivalentés:

(1) T est orthogonale.

(2) $[T]_{\{u_1,\ldots,u_n\}}$ est orthogonale.

(3) $\{T(u_1), \ldots, T(u_n)\}$ est une base orthnormée de E.

Démonstration. On sait que $[T^*] = [T]^T$. Ainsi T est orthogonal si et seulement si $T^* = T^{-1}$ si et seulement si $[T^*] = [T^{-1}]$ si et seulement si $[T]^T = [T]^{-1}$, c'est-à-dire, $[T]$ est orthogonale. Ceci montre l'équivalence de (1) et (2). Enfin, $\{T(u_1), \ldots, T(u_n)\}$ est une base orthnormée si, et seulement si, la matrice de coordonnées de $\{T(u_1), \ldots, T(u_n)\}$ dans $\{u_1, \ldots, u_n\}$ est orthogonale, c'est-à-dire, $[T]_{\{u_1,\ldots,u_n\}}$ est orthogonale. Ceci montre l'équivalence de (2) et (3). La preuve s'achève.

Exemple. (1) La rotation ρ_θ d'angle θ du plan est orthogonale car sa matrice dans la base canonique est

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

qui est orthogonale.

(2) La rotation ρ_θ d'angle θ de $\mathbb{R}^3$ autour l'axe des x est linéaire et telle que $\rho_\theta(e_1) = e_1$, $\rho_\theta(e_2) = (0, \cos\theta, \sin\theta)$, et $\rho_\theta(e_3) = (0, -\sin\theta, \cos\theta)$. On voit que ρ_θ is orthogonale.

9.7. Diagonalisation orthogonale des endomorphismes auto-adjoints

Partout dans cette section, on se fixe E un espace euclidien muni d'un produit scalaire $<\, ,\, >$.

9.7.1. Lemme. Soit $T \in \text{End}(T)$. Si T est auto-adjoint, alors le polynôme caractéristique $\chi_T(\lambda)$ est scindé sur $\mathbb{R}$.

Démonstration. Prenons une base orthonormée $\{u_1, \ldots, u_n\}$ de E. Comme T est auto-adjoint, $[T]_{\{u_1,\ldots,u_n\}}$ est symétrique. Ainsi $\chi_T(\lambda) = \chi_{[T]}(\lambda)$ est scindé sur $\mathbb{R}$. Ceci achève la démonstration.

9.7.2. Lemme. Soit T un endomorphisme auto-adjoint de E. Soient u_1, u_2 des vecteurs propres de T associcés à des valeurs propres λ_1, λ_2 respectivement. Si $\lambda_1 \neq \lambda_2$, alors u_1 et u_2 sont orthogonaux.

Démonstration. On a $T^* = T$ et $T(u_i) = \lambda_i u_i,\ i = 1, 2$. Ainsi

$$\begin{aligned} & \lambda_1 < u_1, u_2 >=< \lambda_1 u_1, u_2 >=< T(u_1), u_2 >=< u_1, T^*(u_2) > \\ = \quad & < u_1, T(u_2) >=< u_1, \lambda_2 u_2 >= \lambda_2 < u_1, u_2 > . \end{aligned}$$

Or $\lambda_1 \neq \lambda_2$ entraîne que $< u_1, u_2 >= 0$. Ceci achève la démonstration.

9.7.3. Théorème. Soit $T \in \text{End}(T)$. Si T est auto-adjoint, alors il existe une base orthonormée de E telle que $[T]_{\{u_1,\ldots,u_n\}}$ est diagonale.

Démonstration. Procédons par induction en $n = \dim(E)$. Si $n = 1$, le résultat est trivial. Supposons $n > 1$ et le résultat est vrai pour tout espace euclidien de dimension $n-1$. D'après la proposition 9.7.1, il existe $u_1 \in E$ non nul tel que $T(u_1) = \lambda_1 u_1,\ \lambda_1 \in \mathbb{R}$. Posons $v_1 = \frac{u_1}{||u_1||}$. Alors v_1 est de longueur 1 et $T(v_1) = \lambda_1 v_1$. Posons F le sous-espace de E engendré par v_1. D'après le théorème 9.4.12, $E = F \oplus F^\perp$ et $\dim(F^\perp) = n-1$. Or pour tout $u \in F^\perp$, on a $< v_1, T(u) >=< v_1, T^*(u) >=< T(v_1), u >=< \lambda_1 v_1, u >= \lambda_1 < v_1, u >= 0$. Ainsi $T(u) \in F^\perp$. Ceci implique que

$$T' : F^\perp \mapsto F^\perp : u \mapsto T(u)$$

est un endomorphisme de $F^\perp$, qui est également auto-adjoint. Par hypothèse de récurrence, il existe une base orthonormée $\{v_2, \ldots, v_n\}$ de $F^\perp$ telle que $[T']_{\{v_2,\ldots,v_n\}} = \text{diag}\{\lambda_2, \ldots, \lambda_n\}$, $\lambda_i \in \mathbb{R}$. C'est-à-dire, $T'(v_i) = t(v_i) = \lambda_i v_i$, pour tout $2 \le i \le n$. D'où $T(v_i) = \lambda_i v_i$, pour tout $1 \le i \le n$. Comme $< v_1, v_i >= 0$, pour tout $2 \le i \le n$, $\{v_1, v_2, \ldots, v_n\}$ est une base orthonormée de E. Or $[T]_{\{v_1,v_2,\ldots,v_n\}} = \text{diag}\,\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Ceci achève la démonstration.

9.7.4. Théorème. Si $A \in M_n(\mathbb{R})$ est symétrique, alors il existe une matrice orthogonale P telle que $P^T AP = P^{-1}AP$ est diagonale.

Démonstration. Considérons l'endomorphisme de l'espace réel $\mathbb{R}^{(n)}$ suivant :

$$T : \mathbb{R}^{(n)} \to \mathbb{R}^{(n)} : u \mapsto Au.$$

Comme A est symétrique, T est auto-adjoint. D'après le théorème 9.7.3, il existe une base orthonormée $\{u_1, \ldots, u_n\}$ de $\mathbb{R}^{(n)}$ telle que $[T]_{\{u_1,\ldots,u_n\}} = \text{diag}\,\{\lambda_1, \ldots, \lambda_n\}$. Remarquons que la matrice de passage de $\{e_1, \ldots, e_n\}$ vers $\{u_1, \ldots, u_n\}$ est $P = (u_1 \cdots u_n)$, ce qui est orthogonale. Ainsi $P^T AP = P^{-1}AP = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$. Ceci achève la démonstration.

9.7.5. Diagonalisation orthogonale de matrices symétriques. Soit $A \in M_n(\mathbb{R})$ symétrique.

(1) Factoriser $\chi_A(\lambda) = (\lambda_1 - \lambda)^{n_1} \cdots (\lambda_r - \lambda)^{n_r}$ avec $n_i > 0$ et $\lambda_1, \ldots, \lambda_r$ deux à deux distincts.

(2) Pour chaque $1 \le i \le r$, trouver une base $\mathcal{B}_i$ de $\mathbb{R}^{(n)}(A, \lambda_i)$.

(3) Pour chaque $1 \le i \le r$, trouver une base orthonormée $\mathcal{O}_i$ de $\mathbb{R}^{(n)}(A, \lambda_i)$ en appliquant le procédé de Gram-Schmidt à $\mathcal{B}_i$. Alors $\mathcal{O} = \mathcal{O}_1 \cup \cdots \cup \mathcal{O}_r$ est une base orthonormée de $\mathbb{R}^{(n)}$.

(4) Poser P la matrice dont les colonnes sont formées des vecteurs de $\mathcal{O}$. Alors P est orthogonale telle que

$$P^T AP = P^{-1}AP = \text{diag}\{\overbrace{\lambda_1, \ldots, \lambda_1}^{n_1}, \ldots, \overbrace{\lambda_r, \ldots, \lambda_r}^{n_r}\}.$$

9.7.6. Diagonalisation orthogonale d'endomorphismes auto-adjoints. Soit T un endomorphisme auto-adjoint de E.

(1) Trouver une base orthonormée $\{u_1, \ldots, u_n\}$ de E et calculer $A = [T]_{\{u_1,\ldots,u_n\}}$.

(2) Trouver une matrice orthogonale P telle que $P^{-1}AP$ est diagonale.

(3) Poser $(v_1 \ldots v_n) = (u_1 \ldots u_n)P$. Alors $\{v_1, \ldots, v_n\}$ est une base orthonormée de E dans laquelle la matrice de T est $P^{-1}AP$.

9.7.7. Théorème. Si q est une forme quadratique sur E, alors il existe une base de E dans laquelle q est de la forme diagonale. En outre, une telle base peut être trouvée de la façon suivante:

(1) Trouver une base quelconque $\{u_1, \ldots, u_n\}$ de E et calculer $A = [q]_{\{u_1,\ldots,u_n\}}$.

(2) Trouver une matrice orthogonale P telle que $P^T AP$ est diagonale.

(3) Poser $(v_1 \ldots v_n) = (u_1 \ldots u_n)P$. Alors $\{v_1, \ldots, v_n\}$ est une base de E dans laquelle la matrice de q est $P^T AP$.

9.8. Exercices

1. Considérer la forme bilinéaire f sur $\mathbb{R}^4$ définie par:

$$f(u,v) = 2x_1y_1 - x_3y_3 + x_4y_4 + 2x_1y_2 - 2x_1y_3 - 2x_3y_1 + 2x_2y_4 - 2x_4y_2 + 5x_3y_4 + 5x_4y_3.$$

   (1) Trouver la matrice de f dans la base canonique de $\mathbb{R}^4$.
   (2) Trouver la matrice de f dans la base $\{(1,1,0,0),\ (1,0,0,1),\ (0,0,1,1),\ (0,1,0,1)\}$.
   (3) Déterminer si f est symétrique ou non.

2. Trouver la forme bilinéaire f sur l'espace réel $\mathbb{R}_3[x]$ dont la matrice dans la base $\{u_1 = 1 + x^2, u_2 = 1 + x, u_3 = x + x^2\}$ est $\text{diag}\{5, -1, -2\}$.

   Attention: Il faut donner la formule explicite de $f(a_0 + a_1x + a_2x^2, b_0 + b_1x + b_2x^2)$.

3. Soit une fonction

$$f : K^{(2)} \times K^{(2)} \to K : (u,v) \mapsto f(u,v).$$

   Montrer que $f(u,v) = \det(uv)$, pour tous $u, v \in K^{(2)}$, si et seulement si f est une forme bilinéaire sur $K^{(2)}$ qui satisfait aux conditions suivantes:

   (1) $f(e_1, e_2) = 1$, où $\{e_1, e_2\}$ est la base canonique de $K^{(2)}$.
   (2) $f(u,u) = 0$, pour tout $u \in K^{(2)}$.
   (3) $f(u,v) = -f(v,u)$, pour tous $u, v \in K^{(2)}$.

4. Dans chacun des cas suivants, déterminer si la forme quadratique est définie positive ou non.

   (1) $q(x_1, x_2, x_3) = x_1^2 + 3x_2^2 + 7x_3^2 + 2x_1x_2 + 8x_2x_3$.
   (2) $q(x_1, x_2, x_3, x_4) = x_1^2 + 2x_2^2 + 2x_1x_2 - 2x_2x_3 - 2x_2x_4 + 4x_3x_4$.

5. Pour quelles valeurs de $\alpha \in \mathbb{R}$ les formes quadratiques sur $\mathbb{R}^3$ ci-dessous sont-elles définies positives?

   (1) $q(x_1, x_2, x_3) = x_1^2 + 6x_2^2 + 3x_3^2 + 4x_1x_2 + 6\alpha x_1x_3$.
   (2) $q(x_1, x_2, x_3) = x_1^2 + 5x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2\alpha x_2x_3$.
   (3) $q(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + x_3^2 + 2\alpha x_1x_2 + 2x_2x_3$.

6. (1) Trouver la forme quadratique q sur $\mathbb{R}_3[x]$ dont la matrice dans la base $\{f_1 = 1+x, f_2 = 1+x^2, f_3 = x + x^2\}$ est $\text{diag}\{2, 3, -1\}$. *Attention*: Il faut donner la formule explicite de $q(a + bx + cx^2)$.

   (2) Montrer que q est définie positive ou non.

7. Soient B, C des matrices carrées réelles et

$$A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}.$$

   Montrer que A est définie positive si et seulement si B, C sont toutes définies positives.

8. Soit $A \in M_n(\mathbb{R})$ définie positive. Montrer, pour toute $B \in M_{n\times m}(\mathbb{R})$, que B^TAB est définie positive si et seulement si B est de rang m.

9. Montrer les énoncés suivants pour $A \in M_n(\mathbb{R})$.

   (1) Si A est définie positive, alors les terms diagonaux de A sont strictement positifs.
   (2) La matrice AA^T est définie positive si et seulement si A est inversible.

10. Si A est une matrice complexe hermitienne, montrer que les mineurs principaux de A sont tous réels.

11. Soit A une matrice réelle anti-symétrique. Montrer les énoncés suivants.

    (1) A a une valeur propre réelle si et seulement si A est non inversible, et dans ce cas, la valeur 0 est la seule valeur propre réelle de A.

    (2) Les matrices $I + A$ et $I - A$ sont toutes inversibles.

12. Pour $u = (x_1, x_2),\ v = (y_1, y_2) \in \mathbb{R}^2$, on pose
$$< u, v >= x_1y_1 - 2x_2y_1 - x_2y_2.$$
Déterminer s'il s'agit d'un produit scalaire sur $\mathbb{R}^2$ ou non.

13. Pour $f = a_0 + a_1x + a_2x^2, g = b_0 + b_1x + b_2x^2 \in \mathbb{R}_2[x]$, on pose
$$< f, g >= (a_0 + a_1)b_0 + (a_0 + 3a_1)b_1 + 3a_2b_2.$$
Montrer qu'il s'agit d'un produit scalaire.

14. Pour tous $u = a_1 + a_2x + a_3x^2, v = b_1 + b_2x + b_3x^2 \in \mathbb{R}_3[x]$, on pose
$$f(u, v) = a_1b_1 + a_1b_2 + \alpha a_1b_3 + a_2b_1 + 2a_2b_2 + a_2b_3 + \alpha a_3b_1 + a_3b_2 + a_3b_3.$$
Donner les valeurs réelles de α pour que f soit un produit scalaire sur $\mathbb{R}_3[x]$.

15. Considérer la forme bilinéaire sur $\mathbb{R}^{(3)}$ suivante:
$$< \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} >= (\frac{1}{\sqrt{3}}x_1 - x_2 + x_3)(\frac{1}{\sqrt{3}}y_1 - y_2 + y_3) + (x_2 - x_3)(y_2 - y_3) + 3x_3y_3.$$
Montrer qu'il s'agit d'un produit scalaire, et trouver l'angle et la distance entre les vecteurs e_1 et e_2 de la base canonique de $\mathbb{R}^{(3)}$ pour ce produit scalaire.

16. Trouver l'angle et la distance entre $\frac{4}{\pi}\sin t$ et $\frac{4}{\pi}\cos t$ dans $C[0, \frac{\pi}{2}]$.

17. Soit E un espace vectoriel réel muni d'un produit scalaire. Montrer le théorème de Pythagore: deux vecteurs u et v de E sont orthogonaux si, et seulement si,
$$||u + v||^2 = ||u||^2 + ||v||^2.$$

18. Donner le produit scalaire $< \, , \, >$ sur l'espace réel $\mathbb{R}_3$ tel que
$$\{u_1 = (1, 1, 0), u_2 = (1, 0, 1), u_3 = (0, 1, 1)\}$$
est une base orthonormée.

19. Soit E un espace vectoriel réel muni d'un produit scalaire avec deux vecteurs non nuls u, v. Montrer que les énoncés suivants.

    (1) L'angle entre u et v est 0 si et seulement si $u = \alpha v$ avec $\alpha > 0$.

    (2) L'angle entre u et v est π si et seulement si $u = \alpha v$ avec $\alpha < 0$.

    Indication: Voir la preuve de l'inégailté de Cauchy-Schwarz.

20. Considérer l'espace vectoriel réel $C[0, 2\pi]$ muni du produit scalaire canonique. Soit $n \geq 1$ un entier.

(1) Montrer que $\{1,\ \sin t,\ \cos t,\ \sin 2t,\ \cos 2t,\ \ldots,\ \sin nt,\ \cos nt\}$ est orthogonale.
(2) En déduire la famille $\{1,\ \sin t, \sin 2t,\ \ldots,\ \sin nt\}$ est libre ou non.
(3) En déduire la famille $\{1,\ \cos t, \cos 2t,\ \ldots,\ \cos nt\}$ est libre ou non.

21. Considérer les fonctions $1,\ t,\ \sin t$ et $\cos t$ dans $C[0, \pi]$.
    (1) Montrer que $\{1,\ t,\ \sin t,\ \cos t\}$ est libre.
    (2) Trouver une base orthogonale du sous-espace de $C[0, \pi]$ engendré par $1,\ t,\ \sin t$ et $\cos t$.
    (3) Trouver la meilleure approximation de e^t par les fonctions $1, t, \sin t$, et $\cos t$.

22. Vérifier que les vecteurs suivants de l'espace réel $\mathbb{R}^{(2)}$ forment une base orthonormée.

$$\left\{ \begin{pmatrix} -\dfrac{1}{\sqrt{5}} \\ \\ \dfrac{2}{\sqrt{5}} \end{pmatrix}, \quad \begin{pmatrix} \dfrac{2}{\sqrt{5}} \\ \\ \dfrac{1}{\sqrt{5}} \end{pmatrix} \right\}.$$

23. Soit F le sous-espace vectoriel de $\mathbb{R}^4$ engendré par les vecteurs suivants:

$$u_1 = (1, 1, 1, 1),\ u_2 = (1, 1, 1, -1),\ u_3 = (1, 1, -1, 1), u_4 = (3, 3, 1, 1).$$

    (1) Trouver une base othonormée de F.
    (2) Trouver la meilleure approximation de $v = (1, 0, 1, 1)$ par des vecteurs de F.

24. Considérer l'espace vectoriel $C[0, 1]$. Soit F le sous-espace engendré par les fonctions $f_1 = 1, f_2 = \sqrt{t}$ et $f_3 = t$. Trouver une base orthogonale de F.

25. À partir la base canonique de $\mathbb{R}^2$, au moyen du procédé de Gram-Schmidt, trouver une base orthonormée de l'espace euclidien $\mathbb{R}^2$ muni du produit scalaire suivant:

$$< (x_1, x_2), (y_1, y_2) >= x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2.$$

26. Considérer l'espace réel $C[0, \pi]$ muni du produit scalaire canonique. Trouver la meilleure approximation de $\cos t$ par des polynômes de degré plus petit ou égal à 3.

27. Soit E un espace euclidien. Si X est un sous-ensemble non vide (pas nécessairement un sous-espace) de E, montrer que

$$X^\perp = \{u \in E \mid\ < x, u >= 0, \text{ pour tout } x \in X\}$$

est un sous-espace de E.

28. On munit $\mathbb{R}^4$ du produit scalaire $<, >$ défini par

$$< u, v >= x_1y_1 + 2x_2y_2 + 4x_3y_3 + 18x_4y_4 + x_1y_3 + x_3y_1 + 2x_2y_4 + 2x_4y_2 + 6x_3y_4 + 6x_4y_3.$$

    (1) Calculer $F^\perp$, où F est l'espace-solution du système suivant:

$$\begin{array}{ccccccccc} x_1 & - & x_2 & + & x_3 & - & x_4 & = & 0 \\ & & x_2 & & & - & 2x_4 & = & 0. \end{array}$$

    (2) Trouver une base orthonormée de $\mathbb{R}^4$ pour ce produit scalaire.

29. Considérer l'espace euclidien $\mathbb{R}^4$ muni du produit scalaire canonique. Soit F le sous-espace de $\mathbb{R}^4$ engendré par $(1,0,1,0)$ et $(0,2,3,1)$. Au moyen du procédé de Gram-Schmidt, trouver une base orthonormée pour F et pour $F^\perp$.

30. Considérer l'espace euclidien $\mathbb{R}_3[x]$ muni du produit scalaire canonique. Trouver l'endomorphisme T de $\mathrm{End}(\mathbb{R}_3[x])$ tel que $T^*(f_1) = f_1 - 2f_2 + f_3$, $T^*(f_2) = 2f_1 - f_2 + f_3$ et $T^*(f_3) = f_2 - f_3$, où $f_1 = 1 + x + x^2, f_2 = x + x^2, f_3 = x - x^2$. *Indication*: Trouver premièrement la matrice de T^* dans la base canonique.

31. Soient E un espace euclidien et T un endomorphisme de E.
   (1) Montrer que $(T^*)^* = T$.
   (2) Montrer que T est orthogonal si et seulement si T^* est orthogonal.
   (3) Si $T^2 = T$, montrer que T est auto-adjoint si et seulement si $TT^* = T^*T$.

32. Trouver des nombres réels a, b, et c tels que la matrice suivante soit orthogonale:

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{14}} & a \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{14}} & b \\ \frac{1}{\sqrt{3}} & -\frac{3}{\sqrt{14}} & c \end{pmatrix}.$$

33. Soit ρ la rotation de $\mathbb{R}^3$ d'angle θ autour de l'axe des y.
   (1) Trouver la matrice de ρ dans la base canonique de $\mathbb{R}^3$.
   (2) Vérifier que ρ est orthogonale.

34. Vérifier l'orthogalité de la réflexion de l'espace euclidien $\mathbb{R}^2$ par rapport à l'axe des x

$$R_x : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (x, -y).$$

35. Dans chacun des cas suivants, trouver la matrice de l'endomorphisme de $\mathbb{R}^3$ dans la base canonique, puis déterminer s'il est orthogonal ou non.
   (1) $T : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \mapsto (\frac{2}{3}x - \frac{1}{3}y + \frac{2}{3}z,\ \frac{2}{3}x + \frac{2}{3}y - \frac{1}{3}z,\ -\frac{1}{3}x + \frac{2}{3}y + \frac{2}{3}z)$.
   (2) $S : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \mapsto (x - 2y + 2z, 2x - 2y + z, -2x + y + 2z)$.

36. Montrer que les matrices orthogonales spéciales d'ordre n forment un groupe pour la multiplication matricielle.

37. Soit $A = (a_{ij})_{n\times n}$ une matrice orthogonale. Montrer les énoncés suivants.
   (1) Si $\det(A) = 1$, alors $c_{ij}(A) = a_{ij}$ pour tous $1 \le i, j \le n$.
   (2) Si $\det(A) = -1$, alors $c_{ij}(A) = -a_{ij}$ pour tous $1 \le i, j \le n$.
   (3) Une valeur propre réelle de A soit 1 soit -1.

38. Soit A une matrice réelle symétrique. Si $A^2 = I$, montrer que $I - 2A$ est orthogonale.

39. Si $u \in \mathbb{R}^{(n)}$ est un vecteur de longueur 1, montrer que $I - 2uu^T$ est une matrice symétrique orthogonale. *Indication*: Utiliser le numéro précédant.

40. Soit A une matrice orthogonale et B une matrice triangulaire supérieure dont les terms diagonaux sont tous positifs. Si AB est aussi une matrice triangulaire supérieure dont les terms diagonaux sont tous positifs, montrer que A est la matrice identité.

    Indication: L'ensemble des matrices triangulaires supérieures de diagonale positive est fermé pour le produit et l'inverse.

41. Soit T un automorphisme orthogonal de E.
    (1) Montrer que $\mathrm{Ker}(T - \mathbf{1}_E) = \mathrm{Im}(T - \mathbf{1}_E)^{\perp}$.
    (2) En déduire que si $(T - \mathbf{1}_E)^2 = 0$, alors $T = \mathbf{1}_E$.

42. Soit E un espace euclidien de dimension impaire. Montrer que si T est un endomorphisme orthogonal de E, alors il existe un vecteur u non nul tel que $T(u) = u$ ou $T(u) = -u$.

43. Diagonaliser orthogonalement chacune des matrices suivantes:

$$(1) \quad \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix} \qquad\qquad (2) \quad \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & -2 & 2 & 1 \\ 1 & 2 & -2 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

44. Diagonaliser chacune des formes quadratiques sur $\mathbb{R}^3$ suivantes:
    (1) $q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$.
    (2) $q(x_1, x_2, x_3) = 5x_1^2 + 52x_2^2 + 2x_3^2 - 8x_1x_2 - 4x_1x_3 + 4x_2x_3$.

45. Considérer l'espace euclidien $\mathbb{C}$ muni du produit scalaire canonique. Soit T l'endomorphise de $\mathbb{C}$ défini par $T(a + bi) = (3a + 2b) + (2a + b)i$. Vérifier que T est auto-adjoint et trouver une base orthonormée de $\mathbb{C}$ dans laquelle la matrice de T est diagonale.

46. Soit $A = (a_{ij})_{n\times n}$ une matrice symétrique réelle avec $\chi_A(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$. Montrer que

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2.$$

    Indication: Considérer la trace de $A^T A$.

47. Soit A une matrice réelle symétrique. Montrer les énoncés suivants.
    (1) A est définie positive si et seulement si toutes ses valeurs propres sont positives.
    (2) Si A est définie positive, alors il existe une matrice réelle B telle que $A = B^2$.
    (3) Si A est définie positive, alors A^{-1} l'est aussi. *Indication*: Utiliser la partie (1).