# Résumé d'Algèbre Linéaire

**Généralités.**

Un *espace vectoriel* sur un corps $K$ est un ensemble $E$ muni d'opérations $E \times E \to E$ (addition, notée $v+w$) et $K \times E \to E$ (multiplication scalaire, noté $\lambda v$) vérifiant certaines règles (axiomes). Ce sont en fait *toutes les règles habituelles* pour ces opérations ($v + w = w + v$, $a(v + w) = av + aw$ etc.). L'ensemble $E$ peut posséder d'autres opérations, mais on les *oublie* en considérant de $E$ comme espace vectoriel.

Les éléments d'un espace vectoriel sur $K$ (ou $K$-espace vectoriel) sont appelés des *vecteurs*, les éléments de $K$ des *scalaires*. Une somme de produits scalaires tel que $c_1u + c_2v + c_3w$ est appelé *combinaison linéaire* des vecteurs utilisés (ici de $u, v, w$), avec *coefficients* les scalaires utilisés (ici de $c_1, c_2, c_3$).

Pour *tout* ensemble $S$, l'ensemble $E = K^S$ des fonctions $S \to K$ avec addition $f+g : s \mapsto f(s)+g(s)$ et multiplication scalaire $\lambda f : s \mapsto \lambda(f(s))$ forme un espace vectoriel sur $K$. On trouve ainsi en particulier les espaces vectoriels $K^n$ (pour $S = \{1, \ldots, n\}$), car $f : \{1, \ldots, n\} \to K$ s'identifie avec le $n$-uplet $(f(1), \ldots, f(n)) \in K^n$), $K^{\mathbf{N}}$ (les suites infinies de scalaires, avec $S = \mathbf{N}$ et une identification similaire), et l'espace $\mathcal{M}_{n,m}$ des matrices $n \times m$ (avec $S = \{1, \ldots, n\} \times \{1, \ldots, m\}$).

Une partie $V \subseteq E$ d'un $K$-espace vectoriel $E$ est dit *sous-espace vectoriel* (s.e.v.) de $E$ si elle vérifie (1) $0 \in V$, (2) $V$ est fermé pour l'addition ($v, w \in V \Rightarrow v + w \in V$), et (3) pour tout scalaire $\lambda$, $V$ est fermé pour la multiplication scalaire par $\lambda$ ($v \in V \Rightarrow \lambda v \in V$). Un s.e.v. $V$ de $E$, muni de l'addition et multiplication scalaire de $E$ (restreints à $V$), est lui-même un $K$-espace vectoriel. Ainsi pour espace vectoriel on prend souvent une partie d'un certain espace $K^S$ dont on montre que c'est un s.e.v., ce qui remplace la vérification de tous les axiomes d'un espace vectoriel par celle d'un s.e.v.. Par exemple l'espace $K[X]$ de polynômes s'identifie au sous-espace $F$ de $K^{\mathbf{N}}$ des suites qui sont nulles à partir d'un certain rang, par $(c_0, c_1, c_2, \ldots) \in F \mapsto c_0 + c_1X + c_2X^2 + \cdots \in K[X]$; l'espace $K[X]_{<d}$ de polynômes de degré inférieur à $d$ est encore un s.e.v. de $K[X]$, etc.).

Une *famille* de vecteurs est une suite (le plus souvent finie) $[v_1, v_2, \ldots]$ de vecteurs $v_i \in E$. Le *sous-espace engendré* par une famille $[v_1, \ldots, v_k]$, noté $\mathrm{Vect}(v_1, \ldots, v_k)$, est l'ensemble des combinaisons linéaires $\{\, c_1v_1 + \cdots + c_kv_k \mid c_1, \ldots, c_k \in K \,\}$ de $v_1, \ldots, v_k$; c'est toujours un s.e.v. de $E$. (Pour une famille infinie $[v_1, v_2, \ldots]$, les éléments de $\mathrm{Vect}(v_1, v_2, \ldots)$ sont des combinaisons linéaires d'un *nombre fini* de vecteurs choisis parmi $v_1, v_2, \ldots$, car une combinaison linéaire ne peut avoir qu'un nombre fini de termes non nuls.) La famille $[v_1, v_2, \ldots]$ est dite *génératrice* d'un s.e.v. donné $V \subseteq E$ si $\mathrm{Vect}(v_1, v_2, \ldots) = V$.

Une famille $[v_1, \ldots, v_k]$ est dite *liée* s'il existe des coefficients $c_1, \ldots, c_k$, dont au moins un coefficient non nul, tels que $c_1v_1+\cdots+c_kv_k = 0$. (Une famille infinie est liée si elle contient un famille finie liée.) Une famille est liée ssi l'un de ces vecteurs s'écrit comme combinaison linéaire des autres. Une *famille libre* est une famille qui n'est pas liée ; $[v_1, \ldots, v_k]$ est libre ssi $c_1v_1 + \cdots + c_kv_k = 0$ entraîne $c_1 = \cdots = c_k = 0$. Une partie d'une famille libre est toujours libre, et la famille vide est toujours libre. Une famille qui étend une famille génératrice de $E$ (c'est-à-dire qui contient une partie génératrice de $E$) est génératrice de $E$.

Une *base* de $V$ est une famille libre et génératrice de $V$. Si $[v_1, \ldots, v_k]$ est une base de $V$ on a $v_1, \ldots, v_k \in V$, et *tout* $v \in V$ s'écrit *de façon unique* comme combinaison linéaire $v = c_1v_1 + \cdots + c_kv_k$ de $v_1, \ldots, v_k$. Ces coefficients (uniques) $c_1, \ldots, c_k$ s'appellent les *coordonnées* de $v$ par rapport à cette base.

**Applications linéaires.**

Une application $f : E \to F$ entre $K$-espaces vectoriels est dite *linéaire* si $f(v + w) = f(v) + f(w)$ et $f(\lambda v) = \lambda f(v)$, ou de façon équivalente $f(\lambda v + \mu w) = \lambda f(v) + \mu f(w)$, pour tout $v, w \in E$ et $\lambda, \mu \in K$. On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires $E \to F$. Tout $f \in \mathcal{L}(E, F)$ vérifie $f(0_E) = 0_F$. La composition d'applications linéaires $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$ donne une application linéaire $g \circ f \in \mathcal{L}(E, G) : v \mapsto g(f(v))$. L'addition $f_1 + f_2 : v \mapsto f_1(v) + f_2(v)$ et la multiplication scalaire $\lambda f : v \mapsto \lambda(f(v))$ donnent à l'ensemble $\mathcal{L}(E, F)$ la structure d'un $K$-espace vectoriel.

Si $f(v) = w$, on appelle $w$ l'image de $v$, et $v$ un antécédent de $w$, pour $f$. L'*image* $f(E)$ de l'application $f$ elle-même (aussi noté $\mathrm{Im}(f)$) est l'ensemble $\{ f(v) \mid v \in E \}$ des vecteurs $w \in F$ qui peuvent s'écrire comme l'image d'au moins un vecteur $v \in E$. Le *noyau* $\ker(f)$ de $f$ est l'ensemble $\{ v \in E \mid f(v) = 0 \}$ des antécédents du vecteur $0 \in F$. On a toujours que $\ker(f)$ est s.e.v. de $E$, et que $\mathrm{Im}(f)$ est s.e.v. de $F$. On a $\ker(f) = \{0\}$ ssi $f$ est injectif, et $\mathrm{Im}(f) = F$ ssi $f$ est surjectif. Si $f$ est bijectif, il existe une application réciproque $f^{-1} : F \to E$ vérifiant $f^{-1} \circ f = \mathrm{id}_E$ et $f \circ f^{-1} = \mathrm{id}_F$, et $f^{-1}$ est, comme $f$, une application *linéaire*. Dans ce cas $f$ est appelé un *isomorphisme* $E \to F$.

Pour tout famille finie $[v_1, \ldots, v_n]$ de vecteurs de $E$, l'application $L : K^n \to E$ formant ses combinaisons linéaires, donnée par $L(c_1, \ldots, c_n) = c_1 v_1 + \cdots + c_k v_k$, est linéaire. La famille est génératrice de $E$ ssi $L$ est surjectif, et la famille est libre ssi $L$ est injectif. Donc si $[v_1, \ldots, v_n]$ est une base $\mathcal{B}$ de $E$, alors $L$ est un *isomorphisme* $K^n \to E$ ; son application réciproque $E \to K^n$ donne l'expression de vecteurs de $E$ en coordonnées par rapport à $\mathcal{B}$. Dans le cas spécial $E = K^n$, la *base canonique* $[\mathbf{e}_1, \ldots, \mathbf{e}_n]$ est telle que pour elle $L = \mathrm{id}_{K^n}$ : on peut écrire $(c_1, \ldots, c_n) = c_1 \mathbf{e}_1 + \cdots + c_n \mathbf{e}_n$ pour tout $(c_1, \ldots, c_n) \in K^n$.

**Dimension.**

Si $E$ admet une famille génératrice finie, on dit que $E$ est de dimension finie. Une famille génératrice de $V$ contient toujours une partie qui reste génératrice de $V$ mais qui est en plus *libre*, autrement dit une base de $V$. En particulier $E$ est de dimension finie ssi $E$ admet une base finie. Une famille libre $[v_1, \ldots, v_k]$ est automatiquement une base de $\mathrm{Vect}(v_1, \ldots, v_k)$. Si $E$ est de dimension finie on peut toujours compléter une telle famille à une base $[v_1, \ldots, v_n]$ (avec $n \geq k$) de $E$ tout entier : le *théorème de la base incomplète.*

Si $E$ admet une base $[v_1, \ldots, v_n]$, alors *toute autre base* de $E$ contient aussi précisément $n$ vecteurs : le théorème de la dimension. Dans ce cas on appelle $n$ la *dimension* de $E$, noté $n = \dim(E)$. Dans ce cas toute famille génératrice contient donc au moins $n$ vecteurs, et toute famille libre au plus $n$ vecteurs. Pour une famille $\mathcal{F}$ de $n$ vecteurs dans un espace de dimension $n$, les conditions "$\mathcal{F}$ est libre" et "$\mathcal{F}$ est génératrice" sont équivalentes, et entraînent donc *chacune* que $\mathcal{F}$ est une base de $E$ (deux conclusions pour le prix d'en vérifier une). Si $\dim(E) = n$ alors tout s.e.v. $V$ de $E$ est de dimension finie, et $0 \leq \dim(V) \leq n$ ; il existe un seul tel $V$ avec $\dim(V) = 0$, à savoir $V = \{0\}$, et un seul tel $V$ avec $\dim(V) = n$, à savoir $V = E$ (mais les dimensions intermédiaires admettent beaucoup de s.e.v. différents).

Tout espace vectoriel $E$ de dimension finie est isomorphe (c'est-à-dire admet une isomorphisme) à un espace vectoriel de la forme $K^n$, à savoir celui avec $n = \dim(E)$. En effet, l'expression en coordonnées par rapport à une base quelconque $[b_1, \ldots, b_n]$ de $E$ est un isomorphisme $E \to K^n$.

**Représentation matricielle.**

Une application linéaire $f \in \mathcal{L}(E, F)$ est entièrement déterminée par ses images $f(v)$ où $v$ parcourt (seulement) une base de $E$. Plus précisément, étant donné une base $\mathcal{E} = [e_1, \ldots, e_m]$ de $E$, l'ensemble $\mathcal{L}(E, F)$ est en *bijection* avec celui des familles $[w_1, \ldots, w_m]$ de $m$ vecteurs $w_j \in F$, en faisant correspondre à $f \in \mathcal{L}(E, F)$ la famille $[f(e_1), \ldots, f(e_m)]$ des images des vecteurs de la base. Les conditions que $f$ soit (a) injectif, (b) surjectif, (c) un isomorphisme sont alors respectivement équivalentes à celles que la famille $[f(e_1), \ldots, f(e_m)]$ soit (a) libre, (b) génératrice de $F$, (c) une base de $F$.

Si en plus de la base $\mathcal{E}$ de $E$ on donne une base $\mathcal{B} = [b_1, \ldots, b_n]$ de $F$, les vecteurs $f(e_1), \ldots, f(e_m)$ peuvent être spécifiés par leur expression en coordonnées sur la base $\mathcal{B}$. On obtient ainsi une représentation matricielle de $f$ : la matrice ${}_{\mathcal{B}}\mathrm{Mat}_{\mathcal{E}}(f) \in \mathcal{M}_{n,m}$ de $f$ par rapport aux bases $\mathcal{E}$ (au départ) et $\mathcal{B}$ (à l'arrivée) contient dans sa colonne $j$ les $n$ coordonnées de $f(e_j) \in F$ par rapport à $\mathcal{B}$, pour $j = 1, \ldots, m$.

**Calcul matriciel.**

On définit les produits matrice-vecteur $\mathcal{M}_{n,m} \times K^m \to K^n$ et matrice-matrice $\mathcal{M}_{n,m} \times \mathcal{M}_{m,l} \to \mathcal{M}_{n,l}$ par les formules connues. Les deux sont compatibles via les isomorphismes $K^n \to \mathcal{M}_{n,1}$ qui identifient $x = (x_1, \ldots, x_n) \in K^n$ avec la matrice $\vec{x}$ à une *colonne* aux coefficients $x_1, \ldots, x_n$ ; on traite donc les vecteurs de $K^n$ comme des "vecteurs colonnes" (même si on écrit les $n$-uplets sur une ligne). Ces produits matriciels sont *associatifs* : $B \cdot (A \cdot \vec{v}) = (B \cdot A) \cdot \vec{v}$ et $C \cdot (B \cdot A) = (C \cdot B) \cdot A$.

Les matrices identité $I_m = (\delta_{i,j})_{i,j=1}^m$ sont élément neutre à gauche et à droite pour le produit matriciel dès qu'il est défini : $I_m \cdot B = B$ pour tout $l$ et $B \in \mathcal{M}_{m,l}$, ainsi que $A \cdot I_m = A$ pour tout $n$ et $A \in \mathcal{M}_{n,m}$. Une matrice $A$ est *inversible* s'il existe une matrice $B$ telle que $B \cdot A$ et $A \cdot B$ sont des matrices identité. Dans ce cas $B$ est unique, et appelé l'inverse $A^{-1}$ de $A$. Toute matrice inversible est carrée. Or, si $A$ est carrée les conditions $B \cdot A = I$ et $A \cdot B = I$ sont équivalentes (deux équations pour le prix d'une).

Si $A = {}_{\mathcal{B}}\operatorname{Mat}_{\mathcal{E}}(f)$ pour $f \in \mathcal{L}(E, F)$ et des bases $\mathcal{E}$ de $E$ et $\mathcal{B}$ de $F$, et si $(c_1, \ldots, c_m) \in K^m$ sont les coordonnées de $v \in E$ par rapport à $\mathcal{E}$, alors $A \cdot \vec{v}$ décrit les coordonnées de $f(v) \in F$ par rapport à $\mathcal{B}$. Le vecteur colonne $A \cdot \vec{v}$ est la combinaison linéaire des colonnes de $A$ avec comme coefficients les coordonnées $c_1, \ldots, c_m$ de $v$, ce qui correspond à $f(v) = f(c_1 e_1 + \cdots + c_n e_n) = c_1 f(e_1) + \cdots + c_m f(e_m)$.

Si $B = {}_{\mathcal{C}}\operatorname{Mat}_{\mathcal{B}}(g)$ et $A = {}_{\mathcal{B}}\operatorname{Mat}_{\mathcal{E}}(f)$, pour $f \in \mathcal{L}(E, F)$, $g \in \mathcal{L}(F, G)$ et des bases $\mathcal{E}, \mathcal{B}, \mathcal{C}$ de respectivement $E, F, G$, alors $B \cdot A = {}_{\mathcal{C}}\operatorname{Mat}_{\mathcal{E}}(g \circ f)$ ; le produit matrice-matrice décrit donc la composition d'applications linéaires, à condition d'utiliser deux fois la même base (ici $\mathcal{B}$) de l'espace vectoriel $F$ au milieu (règle des dominos). L'application $f$ est un isomorphisme ssi sa matrice $A$ est (carrée et) inversible.

Si $\mathcal{E}, \mathcal{B}$ sont deux bases d'un même espace $E$, la matrice $C = {}_{\mathcal{E}}\operatorname{Mat}_{\mathcal{B}}(\mathrm{id}_E)$, dont les colonnes expriment les vecteurs de $\mathcal{B}$ en coordonnées par rapport à $\mathcal{E}$, est appelé *matrice de passage* de $\mathcal{E}$ à $\mathcal{B}$ ; elle est inversible, et $C^{-1} = {}_{\mathcal{B}}\operatorname{Mat}_{\mathcal{E}}(\mathrm{id}_E)$. Si $\vec{v}$ contient les coordonnées de $v$ par rapport à $\mathcal{B}$, alors $C \cdot \vec{v}$ contient ses coordonnées par rapport à $\mathcal{E}$. Les changements de base concernant les matrices découlent de la règle des dominos ; en particulier ${}_{\mathcal{B}}\operatorname{Mat}_{\mathcal{B}}(f) = {}_{\mathcal{B}}\operatorname{Mat}_{\mathcal{E}}(\mathrm{id}_E) \cdot {}_{\mathcal{E}}\operatorname{Mat}_{\mathcal{E}}(f) \cdot {}_{\mathcal{E}}\operatorname{Mat}_{\mathcal{B}}(\mathrm{id}_E) = C^{-1} \cdot {}_{\mathcal{E}}\operatorname{Mat}_{\mathcal{E}}(f) \cdot C$.

Si $A \in \mathcal{M}_{n,m}$, une équation $A \cdot \vec{x} = \vec{b}$ équivaut à un système linéaire de $n$ équations (une pour chaque ligne de $A$ et le coefficient correspondant de $\vec{b}$) en $m$ inconnues (les coefficients de $\vec{x}$). Ainsi des équations linéaires vectorielles, comme la recherche $f(v) = w$ d'un antécédent $v$ de $w$ pour $f$, se résolvent par des méthodes pour les systèmes linéaires. La méthode du pivot de Gauss consiste à appliquer des opérations élémentaires sur les lignes pour échelonner la matrice : dans chaque ligne non nulle, le premier coefficient non nul (le pivot) est aussi le dernier coefficient non nul de sa colonne (en particulier les pivots sont dans des colonnes distinctes). On peut également obtenir un forme échelonnée réduite, où chaque pivot est 1 et est l'*unique* coefficient non nul dans sa colonne. Chaque opération élémentaire équivaut à une multiplication à gauche par une matrice inversible, donc au total on transforme $A \cdot \vec{x} = \vec{b}$ en le système équivalent $G \cdot A \cdot \vec{x} = G \cdot \vec{b}$, où $G$ est un matrice inversible, et $G \cdot A$ échelonnée (éventuellement réduite).

Un système échelonné $A' \cdot \vec{x} = \vec{b}'$ n'a aucune situation s'il existe une ligne nulle de $A'$ pour laquelle le coefficient correspondant de $b'$ n'est pas nul. Sinon une solution générale est obtenue en considérant comme paramètres les inconnues dont la colonne ne contient pas de pivot, et en exprimant chaque autre inconnue en termes de celles-là à l'aide de l'équation contentant son pivot. Cette solution s'écrit aussi comme la somme d'une solution particulière (choisie), et les éléments de $\ker(A) = \ker(A')$ décrit par le système homogène $A' \cdot \vec{x} = \vec{0}$. Or $\dim(\ker(A))$ est égal au nombre de paramètres (colonnes sans pivot), et on en trouve une base en choisissant chaque fois un paramètre égal à 1 et tous les autres égaux à 0.

**Rang.**

Le rang de $f \in \mathcal{L}(E, F)$ est défini par $\operatorname{rg}(f) = \dim(\operatorname{Im}(f))$. Si on forme la matrice $A = {}_{\mathcal{B}}\operatorname{Mat}_{\mathcal{E}}(f)$, pour bases $\mathcal{E}$ de $E$ et $\mathcal{B}$ de $F$, alors $\operatorname{rg}(f) = \operatorname{rg}(A)$, ce rang de $A$ étant défini comme le nombre maximal de colonnes de $A$ qu'on peut choisir qui forment une famille libre. Comme $\operatorname{rg}(f)$ est défini sans référence aux bases, $\operatorname{rg}(A)$ est invariant quand on applique des opérations sur les lignes ou sur les colonnes de $A$. Pour $A \in \mathcal{M}_{n,m}$ on a $0 \leq \operatorname{rg}(A) \leq \min(n, m)$. Pour $A \in \mathcal{M}_{n,n}$ on a $A$ inversible ssi $\operatorname{rg}(A) = n$. Pour une matrice échelonnée, le rang est égal au nombre de pivots (ou de lignes non nulles).

*Théorème du rang*: $\dim(\ker(f)) + \operatorname{rg}(f) = \dim(E)$. (Compter les colonnes de $A'$, sans ou avec pivot.)

**Intersections et sommes de sous-espaces.**

Toute intersection de sous-espace d'une espace vectoriel $E$ est un sous-espace. La somme

$$V_1 + \cdots + V_k \overset{\text{déf}}{=} \{\, v_1 + \cdots + v_k \mid v_1 \in V_1, \ldots, v_k \in V_k \,\}$$

de sous-espaces $V_1, \ldots, V_k$ de $E$ est un sous-espace de $E$. Une réunion de familles génératrices de chacun des $V_i$ est une famille génératrice de $V_1 + \cdots + V_k$ ; on a $\dim(W_1 + \cdots + W_k) \leq \dim(W_1) + \cdots + \dim(W_k)$.

La somme $V_1 + \cdots + V_k$ est appelée une *somme directe* si la réalisation de chacun de ses vecteurs comme $v_1 + \cdots + v_k$ est *unique*, ou de façon équivalente si $v_1 + \cdots + v_k = 0$ entraîne que $v_i = 0$ pour chaque $i$. Pour indiquer que la somme est directe, elle peut être écrite $V_1 \oplus \cdots \oplus V_k$ (ou $\bigoplus_{i=1}^k V_i$). Une réunion de bases de chacun des $V_i$ donne une base de $V_1 \oplus \cdots \oplus V_k$. Donc $\dim(V_1 \oplus \cdots \oplus V_k) = \dim(V_1) + \cdots + \dim(V_k)$, et on a $\dim(V_1 + \cdots + V_k) = \dim(V_1) + \cdots + \dim(V_k)$ *seulement si* la somme $V_1 + \cdots + V_k$ est directe.

Pour le cas de *deux* sous-espaces $V, W$, la somme $V + W$ est directe ssi $V \cap W = \{0\}$. Plus généralement on a

$$\dim(V + W) = \dim(V) + \dim(W) - \dim(V \cap W).$$

Montrer qu'une somme de plusieurs s.e.v. est directe *ne peut pas être fait* en les comparant deux à deux. Cependant, on a

$$(V_1 \oplus \cdots \oplus V_{k-1}) \oplus V_k = V_1 \oplus \cdots \oplus V_k \ :$$

pour vérifier que la somme de $V_1, \ldots, V_k$ est directe il suffit de vérifier (souvent par récurrence) que la somme des $k - 1$ sous-espaces $V_1, \ldots, V_{k-1}$ est directe, et que la somme $(V_1 \oplus \cdots \oplus V_{k-1}) + V_k$ est directe.

**Déterminant.**

Le déterminant $\det(A)$ n'est défini que si $A$ est une matrice carrée. Il est donné par une formule en les coefficients de $A$, qui (bien que très compliquée quand la taille de $A$ est grande) qui n'utilise que les opérations d'addition, de soustraction et de multiplication. Il ne change pas si l'on ajoute un multiple d'une de ses lignes à une autre ligne, ou un multiple d'une de ses colonnes à une autre colonne. Si l'on multiplie par $c$ les coefficients d'une ligne ou d'une colonne, le déterminant est multiplié par $c$. Le déterminant d'une matrice triangulaire est égal au produit des coefficients sur la diagonale; celui d'une matrice triangulaire en blocs est égal au produit des déterminants des blocs (carrés) sur la diagonale.

Une matrice carrée $A$ est inversible ssi $\det(A) \neq 0$. Le déterminant de $A$ est nul s'il possède une ligne ou colonne nulle, deux lignes égales ou deux colonnes égales, ou plus généralement si un ensemble de lignes ou de colonnes est lié. On a $\det(A \cdot B) = \det(A)\det(B)$, et $\det(A^{-1}) = \det(A)^{-1}$ si $\det(A) \neq 0$.

**Endomorphismes, polynômes en un endomorphisme, polynômes annulateurs.**

Une application linéaire $\phi$ dont le domaine et le codomaine sont le même espace vectoriel $E$ est appelé *endomorphisme* de $E$. L'espace $\mathcal{L}(E, E)$ des endomorphismes de $E$ est noté $\mathrm{End}(E)$, ou parfois $\mathcal{L}(E)$. La matrice (carrée) d'un endomorphisme est toujours exprimée sur la *même* base $\mathcal{B}$ au départ et à l'arrivée, et est abrégée $\mathrm{Mat}_{\mathcal{B}}(\phi) \in \mathcal{M}_n(K)$ au lieu de ${}_{\mathcal{B}}\mathrm{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_{n,n}(K)$. Toute matrice $C^{-1}AC$ obtenue à partir d'une matrice carrée $A \in \mathcal{M}_n(K)$ par changement de base (avec matrice de passage $C$) est dite *semblable* à $A$. Pour $\phi \in \mathrm{End}(E)$ et bases $\mathcal{E}, \mathcal{B}$ de $E$, les matrices $\mathrm{Mat}_{\mathcal{E}}(\phi)$ et $\mathrm{Mat}_{\mathcal{B}}(\phi)$ sont semblables.

La composition d'endomorphismes est notée $\psi \circ \phi$ (prononcé $\psi$ après $\phi$), et définit une multiplication non-commutative sur $\mathrm{End}(E)$ (les règles associative et distributives sont valables). Cela permet de définir les puissances $\phi^n$ de $\phi \in \mathrm{End}(E)$ pour $n \in \mathbf{N}$ ; on a $\phi^0 = \mathrm{id}_E$. Les combinaisons linéaires de puissances de $\phi$ donnent les polynômes $\psi = c_0\phi^0 + \cdots + c_k\phi^k \in \mathrm{End}(E)$ en (l'endomorphisme) $\phi$. On note $\psi = P[\phi]$, où $P$ est le polynôme $c_0X^0 + \cdots + c_kX^k \in k[X]$. Si $\mathcal{B}$ est une base quelconque de $E$ et $A = \mathrm{Mat}_{\mathcal{B}}(\phi)$ alors $\mathrm{Mat}_{\mathcal{B}}(P[\phi]) = P[A] = c_0A^0 + \cdots + c_kA^k$ (avec $A^0 = I$). Pour $P, Q \in K[X]$ on a $(P+Q)[\phi] = P[\phi] + Q[\phi]$ ainsi que $(PQ)[\phi] = P[\phi] \circ Q[\phi]$, donc en particulier les polynômes en $\phi$ commutent pour la composition.

Un $P \in K[X]$ tel que $P[\phi] = 0 \in \mathrm{End}(E)$ est appelé *polynôme annulateur* de $\phi$.

Un s.e.v. $V \subseteq E$ tel que $\phi(V) \subseteq V$ est dit $\phi$-stable. Le noyau et l'image d'un polynôme $P[\phi]$ en $\phi$ (et plus généralement d'un endomorphisme qui commute avec $\phi$) sont toujours des sous-espaces $\phi$-stables. Si un produit $PQ$ de polynômes est annulateur de $\phi$, alors $\ker(P[\phi]) \supseteq \mathrm{Im}(Q[\phi])$ et $\ker(Q[\phi]) \supseteq \mathrm{Im}(P[\phi])$.

**Vecteurs propres, valeurs propres, diagonalisation.**

Un *vecteur propre* $v$ de $\phi$ est un vecteur *non nul* tel que $\mathrm{Vect}(v)$ est $\phi$-stable. Alors $\phi(v) = \lambda v$ pour une valeur $\lambda \in K$ dite la *valeur propre* associée au vecteur propre $v$. Les vecteurs propres avec valeur propre $\lambda$ sont les vecteurs non nuls dans $E_\lambda = \ker(\phi - \lambda\,\mathrm{id}_E)$, espace appelé le *sous-espace propre* pour $\lambda$ de $\phi$. Si $\lambda_1, \ldots, \lambda_k$ sont des valeurs *distinctes*, alors la somme $E_{\lambda_1} + \cdots + E_{\lambda_k}$ est toujours *directe*. Le nombre de valeurs propres de $\phi$ ne peut donc pas dépasser $\dim(E)$.

On appelle $\phi$ *diagonalisable* s'il existe une base $\mathcal{C}$ de $E$, dite une *base de diagonalisation* pour $\phi$, pour laquelle $\mathrm{Mat}_\mathcal{C}(\phi)$ est *diagonale*. C'est le cas si et seulement si tous les vecteurs de la base $\mathcal{C}$ sont des vecteurs propres pour $\phi$ ; les coefficients diagonaux de la matrice sont les valeurs propres correspondantes. La condition "$\phi$ diagonalisable" est équivalente à : (i) $\dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) \geq \dim(E)$ pour certains $\lambda_1, \ldots, \lambda_k \in K$ distincts, et (ii) $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} = E$, autrement dit la somme des sous-espaces propres remplit $E$ tout entier (ici chaque valeur propre de $\phi$ est présente une seule fois parmi $\lambda_1, \ldots, \lambda_k \in K$). Alors les bases de diagonalisation pour $\phi$ sont celles obtenues comme réunion de bases choisies dans chacun des sous-espaces propres $E_{\lambda_i}$. Une matrice $A \in \mathcal{M}_n(K)$ est diagonalisable se elle est semblable à une matrice diagonale ; donc si $\phi$ est diagonalisable, toute matrice $\mathrm{Mat}_\mathcal{B}(\phi)$ est diagonalisable.

Un vecteur propre de $\phi$ pour la valeur propre $\lambda$ est aussi vecteur propre de $P[\phi]$, pour la valeur propre $P[\lambda]$, quel que soit $P \in K[X]$. Changement de base vers une base de diagonalisation (si elle existe) facilite le calcul des puissances de $A \in \mathcal{M}_n(K)$ : si $D = C^{-1}AC$ est diagonale avec coefficients diagonaux $\lambda_1, \ldots, \lambda_n$, alors la matrice $D^k$ est diagonale avec coefficients diagonaux $\lambda_1^k, \ldots, \lambda_n^k$ et $A^k = CD^kC^{-1}$ ; plus généralement $P[A] = C\,P[D]\,C^{-1}$ pour tout $P \in K[X]$, et la matrice $P[D]$ est diagonale avec coefficients diagonaux $P[\lambda_1], \ldots, P[\lambda_n]$. Si $\phi$ est diagonalisable, alors les polynômes annulateurs de $\phi$ sont donc précisément les $P \in K[X]$ tels que $P[\lambda] = 0$ pour toute valeur propre $\lambda$ de $\phi$.

**Polynôme caractéristique.**

Le *polynôme caractéristique* de $A \in \mathcal{M}_n(K)$ est $\chi_A = \det(XI_n - A) \in K[X]$, un polynôme unitaire de degré $n$. Des matrices semblables ont le même polynôme caractéristique, ce qui permet de définir le polynôme caractéristique $\chi_\phi$ de $\phi \in \mathrm{End}(E)$ comme celui de $\mathrm{Mat}_\mathcal{B}[\phi]$ pour n'importe quelle base $\mathcal{B}$ de $E$. Les valeurs propres de $\phi$ sont précisément les racines de $\chi_\phi$. Si $\lambda$ est racine de $\chi_\phi$ avec multiplicité $m$, alors $1 \leq \dim(E_\lambda) \leq m$. Dans ce cas $\phi$ ne peut être diagonalisable que si en fait $\dim(E_\lambda) = m$. Une condition nécessaire et suffisante pour que $\phi$ soit diagonalisable est qu'on puisse décomposer $\chi_\phi = (X - \lambda_1)^{m_1} \ldots (X - \lambda_k)^{m_k}$ avec $\lambda_1, \ldots, \lambda_k \in K$ distincts, et qu'on ait $\dim(E_{\lambda_i}) = m_i$ pour tout $m_i > 1$.

**Polynômes en $X$ à coefficients dans le corps $K$.**

Dans $K[X]$ on a une division euclidienne : pour $P, B \in K[X]$ avec $B \neq 0$ il existe $Q, R \in K[X]$ (dits respectivement *quotient* et *reste* dans la division euclidienne de $P$ par $B$) tels que $P = BQ + R$ et $\deg R < \deg B$. Le couple $(Q, R)$ est unique. Si $R = 0$, on dit que la division est exacte, ou que $B$ divise $P$ ou que $P$ est divisible par $B$ ; dans ce cas on écrit $Q = P/B$. Dans le cas particulier $B = X - c$, le reste est de degré $< 1$ donc un polynôme constant ; sa valeur est l'évaluation $P[c]$ de (la fonction polynomiale de) $P$ en $c$. En particulier $P$ est divisible par $X - c$ ssi $c$ est *racine* de $P$ (c'est-à-dire, si $P[c] = 0$). La *multiplicité* de $c$ comme racine de $P \neq 0$ est le plus grand $m \in \mathbf{N}$ tel que $(X - c)^m$ divise $P$.

Le *plus petit multiple commun* (noté ppcm) de deux (ou plusieurs) polynômes est le multiple commun unitaire de plus petit degré possible ; il divise tous leurs multiples communs. Le *plus grand diviseur commun* de $A, B \in K[X]$ (noté $\mathrm{pgcd}(A, B)$) est le polynôme unitaire du plus petit degré possible qui soit une combinaison polynomiale de $A$ et $B$ : $\mathrm{pgcd}(A, B) = SA + TB$ (relation de Bézout) pour certains $S, T \in K[X]$ (appelés *coefficients de Bézout* pour $A, B$). On a que $\mathrm{pgcd}(A, B)$ divise toutes ces combinaisons polynomiales de $A, B$ ; en particulier c'est (comme le nomme le suggère) un diviseur commun de $A, B$. Aussi tout diviseur commun de $A, B$ divise $\mathrm{pgcd}(A, B)$ (et $\mathrm{pgcd}(A, B)$ est donc du plus grand degré possible parmi les diviseurs communs de $A, B$).

Le *polynôme minimal* de $\phi \in \mathrm{End}(E)$ est un polynôme annulateur de $\phi$, unitaire et de plus petit degré possible ; il divise tous les polynômes annulateurs de $\phi$.

**Factorisation dans $K[X]$.**

Dans $K[X]$ on a une théorie de factorisation analogue à celle dans $\mathbf{Z}$. Un polynôme non constant est réductible s'il se décompose en produit de deux facteurs de degré $> 0$, sinon irréductible (notions analogues aux entiers composés et premiers; un polynôme constant n'est ni réductible ni irréductible). Les polynômes de degré 1 sont tous irréductibles. Si $K = \mathbf{C}$ il n'y a pas d'autres polynômes irréductibles, mais $K = \mathbf{R}$ en connaît de degré 2 (de discriminant $< 0$) ; pour $K = \mathbf{Q}$, tous les degrés $> 0$ sont possibles.

Tout polynôme $P \neq 0$ s'écrit comme le produit d'une constante (son coefficient dominant) et d'un nombre (fini) de polynômes unitaires irréductibles (un même facteur peut être utilisé plusieurs fois). Cette écriture est *unique* à l'ordre des facteurs irréductibles près ; on l'appelle la *factorisation* de $P$ dans $K[X]$.

Deux polynômes $A, B$ sont *premiers entre eux* si $\operatorname{pgcd}(A, B) = 1$. Une relation de Bézout prend alors la forme $1 = SA + TB$. Deux produits $A_1 \cdots A_k$ et $B_1 \cdots B_l$ sont premiers entre eux si et seulement si $A_i$ et $B_j$ sont premiers entre eux pour tout $1 \leq i \leq k$ et tout $1 \leq j \leq l$.

**Théorème de décomposition des noyaux.**

Si l'on décompose $P = P_1 \cdots P_k$ comme produit de polynômes $P_i$ qui sont premiers entre eux 2 à 2, alors pour tout $\phi \in \operatorname{End}(E)$ on a la décomposition en somme directe

$$\ker(P[\phi]) = \ker(P_1[\phi]) \oplus \cdots \oplus \ker(P_k[\phi]),$$

le *théorème de décomposition des noyaux.* En particulier, si le produit $P$ est un polynôme annulateur de $\phi$ (par exemple si $P = \mu_\phi$), cette somme remplit l'espace vectoriel : $E = \ker(P_1[\phi]) \oplus \cdots \oplus \ker(P_k[\phi])$.

**Réduction d'endomorphismes.**

Le polynôme minimal $\mu_\phi$ de $\phi \in \operatorname{End}(E)$ a pour racines précisément les valeurs propres de $\phi$. Il est scindé et a racines simples si et seulement si $\phi$ est diagonalisable, et cette condition est déjà assurée si $\phi$ possède un polynôme annulateur qui est scindé et à racines simples.

La restriction $\phi|_V$ de $\phi$ à un sous-espace $\phi$-stable $V$ a un polynôme minimal qui est une diviseur unitaire du polynôme minimal $\mu_\phi$. Et si $P$ est un diviseur unitaire quelconque de $\mu_\phi$, alors le quotient $Q = \mu_\phi / P$ est le polynôme minimal de la restriction $\phi|_W$ à l'image $W = \operatorname{Im}(P[\phi])$. La restriction de $\phi$ à $\operatorname{Ker}(Q[\phi])$ a également $Q$ comme polynôme minimal, pour tout diviseur unitaire $Q$ de $\mu_\phi$.

La «réduction d'un endomorphisme $\phi$» est donnée par le théorème de décomposition des noyaux, appliqué pour une décomposition de $\mu_\phi$. Si $\mu_\phi$ est scindé, on prend la décomposition $\mu_\phi = P_1 \cdots P_k$ avec $P_i = (X - \lambda_i)^{m_i}$ pour des valeurs propres $\lambda_i$ distinctes. Le *sous-espace caractéristique* $\tilde{E}_{\lambda_i}$ pour $\lambda_i$ est alors donné par

$$\tilde{E}_{\lambda_i} = \ker(P_i[\phi])) = \ker((\phi - \lambda_i \operatorname{id}_E)^{m_i}),$$

et $E = \bigoplus_i \tilde{E}_{\lambda_i}$ décompose l'espace $E$ comme la somme directe des sous-espaces caractéristiques pour les valeurs propres $\lambda_i$. Le sous-espace caractéristique $\tilde{E}_{\lambda_i}$ contient l'espace propre $E_{\lambda_i} = \ker(\phi - \lambda \operatorname{id})$. En fait, il contient une chaîne *strictement* croissante de sous-espaces

$$\{0\} \subset E_{\lambda_i} = \ker(\psi) \subset \ker(\psi^2) \subset \cdots \subset \ker(\psi^{m_i}) = \tilde{E}_{\lambda_i} \qquad \text{où } \psi = \phi - \lambda_i \operatorname{id}_E,$$

et pour des exposants au delà de $m_i$ le noyau ne change plus : $\ker(\psi^l) = \tilde{E}_{\lambda_i}$ pour tout $l \geq m_i$. Par conséquent $m_i \leq \dim(E_{\lambda_i})$ ; cette dimension est la multiplicité de $\lambda_i$ comme racine du polynôme $\chi_\phi$.

Pour une base adaptée a cette chaîne se sous-espaces, la matrice de la restriction de $\phi$ à $\tilde{E}_{\lambda_i}$ est triangulaire supérieure avec coefficients diagonaux tous $\lambda_i$. Alors $\phi$ est *trigonalisable* (admet une base pour laquelle sa matrice est triangulaire supérieure) si et seulement si $\mu_\phi$ est scindé, car c'est (uniquement) dans ce cas qu'on a une décomposition de $E$ en sous-espaces caractéristiques.

**Rapport entre polynômes minimal et caractéristique.**

La multiplicité de $\lambda_i$ comme racine du polynôme caractéristique $\chi_\lambda$ est égale à $\dim(E_{\lambda_i})$. Le polynôme minimal $\mu_\phi$ divise toujours le polynôme caractéristique $\chi_\phi$, autrement dit, $\chi_\lambda$ est un polynôme annulateur de $\phi$ : c'est le *théorème de Cayley–Hamilton.* Dans le sens opposé, chaque facteur irréductible de $\chi_\phi$ divise $\mu_\phi$ (peut-être avec une plus petite multiplicité). On a $\mu_\phi$ scindé si et seulement si $\chi_\phi$ est scindé.