
Révisions d'algèbre linéaire

Dans cette note, je vais passer en revue, de façon plus ou moins succincte, les notions de l'UE Algèbre II qui me paraissent nécessaires pour une bonne compréhension de l'Algèbre III (et du IV).

1. Espace, sous-espace vectoriel

Un espace vectoriel est la donnée d'un ensemble non vide E et d'un corps commutatif $\mathbb{K}$ (souvent $\mathbb{R}$ ou $\mathbb{C}$) muni de deux lois, autrement dit deux applications :

— L'addition interne : $E \times E \to E,\ (x, y) \mapsto x + y$

— La multiplication externe : $\mathbb{K} \times E \to E,\ (\lambda, x) \mapsto \lambda \cdot x$,

ces lois satisfaisant les axiomes suivants.

1. L'addition est commutative, associative, possède un élément neutre (qu'on note 0_E ou $\vec{0}$ ou simplement 0) et tout élément de E possède un inverse pour cette addition.
2. Pour tous $x, y \in E$ et $\lambda, \mu \in \mathbb{K}$, on a :

$$(\lambda +_{\mathbb{K}} \mu) \cdot x = \lambda \cdot x +_E \mu \cdot x, \quad \lambda \cdot (x +_E y) = \lambda \cdot x +_E \lambda \cdot y, \quad (\lambda\mu) \cdot x = \lambda \cdot (\mu \cdot x), \quad 1_{\mathbb{K}} \cdot x = x.$$

Les exemples les plus courants d'espaces vectoriels sont : $\mathbb{R}$, $\mathbb{C}$ (en tant que $\mathbb{R}$-esp. vect ou $\mathbb{C}$-esp. vect.), $\mathbb{R}^n$, $\mathbb{C}^n$, $M_{n\times p}(\mathbb{K})$, $\mathbb{K}[x]$, $\mathbb{K}_n[x]$. Il y a aussi l'espace vectoriel des suites réelles ou complexes, l'espace des fonctions d'un intervalle de $\mathbb{R}$ vers $\mathbb{R}$, etc.
Un sous-espace vectoriel est un sous-ensemble non vide F d'un espace vectoriel E stable par les deux lois de E (i.e. si $x, y \in F$ alors $x + y \in F$ et si $\lambda \in \mathbb{K}$ et $x \in F$ alors $\lambda \cdot x \in F$). Un sous-espace vectoriel est alors lui-même un espace vectoriel.

En pratique, pour montrer qu'une partie F d'un espace vectoriel E est un sous-espace vectoriel, on commence par vérifier que le zéro de E appartient bien à F (c'est une condition nécessaire et justifie le fait que F n'est pas vide) puis on fait comme suit :

Soient $x, y \in F$ et $\lambda \in \mathbb{K}$.
...
... donc $\lambda x + y \in F$.

2. Famille libre, génératrice. Base et dimension

Soit E un $\mathbb{K}$-espace vectoriel. Une famille finie $v_1, \ldots, v_m$ est dite libre si pour tous $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$, l'égalité $\sum \lambda_i v_i = 0$ implique $\lambda_1 = \cdots = \lambda_m = 0$. Pour montrer qu'une telle famille est libre, en général on rédige de la façon suivante.

Soient $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ tels que $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$.
...
... donc les λ_i sont tous nuls. La famille des v_i est donc libre.

Étant donnée une famille non finie, on dit qu'elle est libre si toute sous-famille finie est libre au sens précédent. N'oublions pas qu'une combinaison linéaire est nécessairement finie.

Une famille (finie) $v_1, \ldots, v_m$ est génératrice de E (on dit aussi que les v_i engendrent E) si tout élément de E est combinaison linéaire des v_i. On peut aussi noter $E = \text{Vect}\{v_1, \ldots, v_m\}$.
Par exemple, soit F un sous-espace vectoriel d'un $\mathbb{K}$-espace vectoriel E et on aimerait montrer qu'une famille $v_1, \ldots, v_m$ engendre F. On doit d'abord s'assurer que les v_i sont bien dans F puis on prend un élément de F ("Soit $x \in F$.") et on montre qu'il existe $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ tels que $x = \lambda_1 v_1 + \cdots + \lambda_m v_m$. Parfois on résout l'équation $x = \lambda_1 v_1 + \cdots + \lambda_m v_m$ dont les inconnues sont les λ_i après avoir posé $x \in F$. On peut aussi utiliser la dimension de F si on la connait et montrer que la famille est libre avant de montrer que c'est une base.

Rappelons qu'une base est une famille qui est libre et génératrice.
Toutes les bases ont le même cardinal et ce cardinal s'appelle la dimension de l'espace en question.
Il faut aussi connaitre le théorème de la base incomplète qui dit deux choses : (a) toute famille libre peut être complétée en une base ; (b) de toute famille génératrice, on peut extraire une base.
Si $v_1, \ldots, v_n$ est une base de E alors pour tout $x \in E$, il existe un unique n-uplet $(x_1, \ldots, x_n) \in \mathbb{K}^n$ tel que $x = \sum_1^n x_i v_i$. Les x_i sont appelées coordonnées de x dans la base en question. Le vecteur $X \in M_{n\times 1}(\mathbb{K})$ (matrice colonne) est appelé vecteur des coordonnées de x dans la base des v_i ou encore matrice (colonne) de x dans la base des v_i.

Attention au terme base canonique. Il ne s'emploie que dans des cas bien précis. Dans l'espace $\mathbb{K}^n$, la base canonique désigne la base suivante : $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$.
Dans l'espace des polynômes $\mathbb{K}[x]$, la base canonique désigne la base constituée de : $1, x, x^2, \ldots, x^k, \ldots$ (la base est infinie).
Dans l'espace $\mathbb{K}_n[x]$, la base canonique est constituée de $1, x, x^2, \ldots, x^n$. La dimension de $\mathbb{K}_n[x]$ est donc $n + 1$.
Dans l'espace des matrices $M_{n\times m}(\mathbb{K})$, la base canonique est la base dont les éléments sont les matrices E_{ij} avec $(i, j) \in \{1, \ldots n\} \times \{1, \ldots, m\}$; E_{ij} étant la matrice ayant un 1 en position (i, j) et des 0 partout ailleurs. Ainsi $M_{n\times m}(\mathbb{K})$ est de dimension nm.
Dans tout autre espace vectoriel que ceux-ci, la notion de base canonique n'a aucun sens.

3. Applications linéaires et matrices

Soient E et E' deux $\mathbb{K}$-espaces vectoriels.
Une application linéaire u de E vers E' est une application telle que pour tous $x, y \in E$ et $\lambda \in \mathbb{K}$, $u(x + y) = u(x) + u(y)$ et $u(\lambda \cdot x) = \lambda \cdot u(x)$. Une application linéaire envoie toujours le vecteur nul de E sur celui de E'.
Rappelons l'image de u : $\text{Im}(u) = \{u(x)\ ;\ x \in E\} = \{x' \in E' \mid \exists x \in E, u(x) = x'\}$; et le noyau de u : $\ker(u) = \{x \in E \mid u(x) = 0\}$.
Le noyau est un sous-esp. vect. de l'ensemble de départ E. L'image est un sous-esp. vect. de l'ensemble d'arrivée E'. On rappelle que u est injective si et s. si $\ker(u) = \{0\}$ (et pas vide !) et que u est surjective si et s. si $\text{Im}(u) = E'$.
Quand on cherche à montrer qu'une application est injective, on procède généralement de la façon suivante.

Soit $x \in E$ tel que $u(x) = 0$ (ou tout simplement "Soit $x \in \ker(u)$").
...
... donc x est nul et donc u est injective.

Il n'est pas utile et encore moins nécessaire de montrer que 0 appartient à $\ker(u)$ puisque le noyau, étant un sous-espace vectoriel, contient toujours le vecteur nul.

En pratique, il est généralement plus fastidieux de montrer qu'une application linéaire est surjective et on contourne souvent le problème en utilisant la dimension des espaces et le théorème du rang (valable en dimensions finies) ; ce théorème dit la chose suivante : la dimension de l'espace de départ (ici E) est égale à la somme des dimensions du noyau et de l'image de u.
Ce théorème nous donne un corollaire utile : si E et E' ont la même dimension finie alors u est injective ssi elle est surjective ssi elle est bijective.

Plaçons nous en dimension finie et fixons des bases. Soit $B = (b_1, \dots, b_n)$ une base de E et soit $B' = (b'_1, \dots, b'_m)$ une base de E'.
La matrice de u relativement aux bases B et B', notée $M_{B,B'}(u)$ est la matrice de $M_{m\times n}(\mathbb{K})$ telle que les colonnes sont les coordonnées de $u(b_1), \dots, u(b_n)$ exprimées dans la base B'.
Soit $x \in E$. Soit $X \in M_{n\times 1}(\mathbb{K})$ la matrice colonne de x dans la base B (dont les termes sont les coordonnées de x). Soit $y = u(x)$ l'image de x par u. Soit Y la matrice colonne de y dans la base B'. Soit $A = M_{B,B'}(u)$. Alors on la relation suivante :

$$A \cdot X = Y.$$

On définit le rang d'une matrice A comme étant la dimension de l'espace vectoriel engendré par les colonnes. Le noyau de A est défini comme $\{X \in M_{n\times 1}(\mathbb{K}) \mid A \cdot X = 0\}$.
En ce qui nous concerne, $\ker(u)$ et $\ker(A)$ sont deux objets de nature différente : l'un est un sev de E et l'autre un sev de $M_{n\times 1}(\mathbb{K})$. Cependant ils sont liés par le fait qu'ils ont la même dimension. De même $\mathrm{rg}(u) = \mathrm{rg}(A)$.
De façon plus précise, si on trouve une base $X_1, \dots, X_p$ de $\ker(A)$ (par la résolution d'un système par exemple). Alors les vecteurs $x_1, \dots, x_p$ de E dont les matrices dans la base B sont les X_i forment une base de $\ker(u)$.
Rappelons ce qui concerne la composée de deux applications linéaires. Soient $u : E \to E'$ et $u' : E' \to E''$. Soient B, B', B'' des bases respectives de E, E' et E''. Soient $A = M_{B,B'}(u)$, $A' = M_{B',B''}(u')$ et $C = M_{B,B''}(u' \circ u)$ alors on a : $C = A' \cdot A$.

4. Changement de bases

Dans la suite, on va se consacrer au cas des endomorphismes (i.e. l'espace d'arrivée est l'espace de départ), je dirai quelques mots plus loin pour expliquer la raison pour laquelle dans le cours d'algèbre III, on ne s'intéresse pas aux applications linéaires générales.
Soit donc $u : E \to E$ un endomorphisme linéaire avec E de dimension finie n sur $\mathbb{K}$.
Soit $B = (b_1, \dots, b_n)$ une base de E et soit $A = M_B(u)$ la matrice de u dans la base B ; on pourrait aussi écrire $M_{B,B}(u)$.
Donnons nous une autre base B' de E et soit $A' = M_{B'}(u)$. Il faut voir B comme l'ancienne base et B' la nouvelle base.

Il y a un lien entre A et A' et ce lien est donné par une matrice de passage ou matrice de changement de base. Soit $P = \text{Pass}_{B,B'}$ la matrice de passage de B vers B'. Par définition, ses colonnes sont les coordonnées des b'_i dans la base B (nouvelle base exprimée dans l'ancienne). On peut aussi définir P comme $M_{B',B}(\text{Id})$.
Cette matrice P est inversible et on a la relation suivante : $P^{-1} \cdot A \cdot P = A'$.

Voici deux définitions (peut-être nouvelles). Soient A et A' deux matrices dans $M_n(\mathbb{K})$.
A et A' sont dites semblables si il existe $P \in M_n(\mathbb{K})$ inversible telle que $P^{-1}AP = A'$.
A et A' sont dites équivalentes s'il existe $P, Q \in M_n(\mathbb{K})$ inversibles telles que $Q^{-1}AP = A'$.

Ainsi A et A' sont semblables si et s. si elles représentent un même endomorphisme dans deux bases différentes, i.e. il existe un endomorphisme $u : \mathbb{K}^n \to \mathbb{K}^n$ et il existe deux bases B, B' de $\mathbb{K}^n$ tels que $A = M_B(u)$ et $A' = M_{B'}(u)$.
Par contre, A et A' sont équivalentes si et s. si elles représentent un même endomorphisme dans deux paires de bases différentes autrement dit il existe un endomorphisme $u : \mathbb{K}^n \to \mathbb{K}^n$ et il existe des bases B_1, B_2, B'_1, B'_2 tels que $A = M_{B_1,B_2}(u)$ et $A' = M_{B'_1,B'_2}(u)$.
En effet, dans la première relation ci-dessus, il faut voir P comme la matrice de passe de B à B' et P^{-1} comme celle de B' à B ; alors que dans la deuxième relation, il faut voir P comme la matrice de passage de B_1 à B'_1 et Q comme la matrice de passage de B_2 à B'_2.
Bien entendu, si $B_1 = B_2$ et $B'_1 = B'_2$ on obtient $Q = P$.
On voit donc que "semblables" implique "équivalentes" mais la réciproque n'est pas du tout vraie en général.
Il y a un théorème (qu'on admet) qui dit que deux matrices sont équivalentes si et s. si elles ont le même rang. La relation d'équivalence n'est pas assez restrictive pour distinguer des endomorphismes qui ont peu de liens géométriques.

Par exemple, soit $u : \mathbb{R}^2 \to \mathbb{R}^2$ l'application définie par $u(x, y) = (-x, -y)$. D'un point de vue géométrique, il s'agit simplement de la symétrie par rapport à l'origine. Sa matrice dans la base canonique de $\mathbb{R}^2$ est $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. Cette matrice est de rang 2, tout comme la matrice identité, donc A est équivalente à la matrice identité pourtant l'application u n'a rien à voir avec l'application identité et en effet la matrice A et la matrice identité ne sont pas semblables (admis pour le moment).

Dans ce cours, le but est d'étudier des endomorphismes en dimension finie et de chercher une base dans laquelle leur matrice aura une forme "simple". On parlera de réduction d'un endomorphisme. Aussi, nous serons amenés à travailler avec des matrices semblables mais nous ne nous occuperons pas de matrices équivalentes.

5. Somme directe et supplémentaires

Une façon d'étudier la réduction des endomrophismes est de "diviser" l'espace E en petits sous-espaces vectoriels dans lesquels on pourra plus facilement étudier l'endomorphisme. Une notion est alors importante, c'est celle de somme directe. Une partie de ce qui suit est nouvelle...
Soit E un espace vectoriel. Soient F et G deux sev de E.

On définit la somme de F et G et on la note $F+G$, comme étant $F+G=\{f+g\ ;\ f\in F, g\in G\}$.
C'est donc l'ensemble de toutes les sommes possibles d'un élément de F et d'un élément de G.
Cette définition s'étend à plus de deux sous-espaces. Si $F_1,\dots,F_m$ sont des sev de E, on définit de la même façon $F_1+\cdots+F_m=\{f_1+\cdots+f_m\ ;\ \forall i=1,\dots,m,\ f_i\in F_i\}$.
Cette somme est elle-même un sev de E.
On dit alors que E est la somme des F_i si $E=F_1+\cdots+F_m$.

Revenons au cas de deux sev F et G.
On dit que E est la somme directe de F et G si $E=F+G$ et $F\cap G=\{0\}$; on écrit alors $E=F\oplus G$.
On dit aussi que F et G sont en somme directe si $F\cap G=\{0\}$.
Ainsi $E=F\oplus G$ signifie que E est la somme de F et G et que ces derniers sont en somme directe.
Exemple : dans $\mathbb{R}^3$, notons (e_1,e_2,e_3) la base canonique. Soit $F=\text{Vect}\{e_1,e_2\}$ et $G=\text{Vect}\{e_2,e_3\}$. Ce sont deux plans qui s'intersectent le long de la droite $\text{Vect}\{e_2\}$. On a bien $E=F+G$ mais $F\cap G\neq\{0\}$.
Autre exemple toujours dans $\mathbb{R}^3$. Soit $F=\text{Vect}\{e_1,e_2\}$ et $G=\text{Vect}\{e_3\}$. Dans ce cas là, on a bien $E=F\oplus G$.

La définition de somme directe s'étend également à plusieurs sev. On dit que E est la somme directe de $F_1,\dots,F_m$, et on le note $E=F_1\oplus\cdots\oplus F_m$ si, d'une part $E=F_1+\cdots+F_m$ et d'autre part : pour tout $(f_1,\dots,f_m)\in F_1\times\cdots\times F_m$, si $f_1+\cdots+f_m=0$ alors tous les f_i sont nuls.
Voici une définition équivalente : $E=F_1\oplus\cdots\oplus F_m$ si tout élément de E s'écrit de façon unique comme une somme de la forme $f_1+\cdots+f_m$ avec $f_i\in F_i$ pour tout i.

Comme précédent, on dit que les F_i sont en somme directe si : pour tout $(f_1,\dots,f_m)\in F_1\times\cdots\times F_m$, l'égalité $f_1+\cdots+f_m=0$ implique que tous les f_i sont nuls.

Il y a une similitude entre famille génératrice et somme ; et entre famille libre et être en somme directe.
En effet, soit $v_1,\dots,v_p$ une famille de vecteurs dans E. Soient $F_i=\text{Vect}\{v_i\}$ pour tout $i=1,\dots,p$.
Alors la famille des v_i est libre si et s. si les F_i sont en somme directe ; et la famille des v_i engendre E si et s. si E est la somme des F_i.

Attention, il est faux de penser que $E=F\oplus G\oplus H$ si et s. si $E=F+G+H$ et $F\cap G\cap H=\{0\}$; ou même $F\cap G=\{0\}$ et $F\cap H=\{0\}$ et $G\cap H=\{0\}$ (voir le TD pour un exemple). Ainsi, la première définition n'est valable que pour deux sev.

6. Révisions/Conseils sur la façon de faire une démonstration.

1. Les assertions du genre : pour tout $x\in X$, etc.
   Apprendre à démarrer par : Soit $x\in X$.
2. Faire attention à différencier "Soit $x\in X$. Alors x est truc" et "Pour tout $x\in X$, x est truc".
3. Dans une démonstration, on n'utilise pas le symbole "$\Rightarrow$" à la place de "donc".
4. On évite au maximum d'utiliser le symbole "$\Longleftrightarrow$".
5. Quand on veut démontrer une assertion du genre : Montrer que A implique B.

(a) On traduit mathématiquement le but B (souvent au brouillon).

(b) On traduit (éventuellement) l'hypothèse A (souvent au brouillon).

(c) On commence la démonstration de B et en cours de route on utilise A.

6. Ce qu'il ne faut pas faire : partir de A pour aller à B.

Exemple 1 (Exemple illustrant le point 3)**.** Dans une démonstration on peut éventuellement utiliser le symbole $\Rightarrow$ mais dans des situations bien précises. Par exemple

Nous allons montrer que l'implication suivante est fausse : $x = 5 \Rightarrow x^2 = 2$.

Dans cet exemple, c'est normal d'utiliser le symbole car on décrit une implication.
Voici un exemple où l'utilisation de $\Rightarrow$ est dangereuse.

On suppose $x > 0$. Alors $y > 3 \Rightarrow z < 0$.

À cause du symbole $\Rightarrow$ cette phrase peut se comprendre de deux façons différentes.
Première interprétation : On suppose $x > 0$. Par conséquent $y > 3$ et cela entraine $z < 0$.
Dans cette première interprétation, la conclusion "$z < 0$" est vraie et est la conséquence de $y > 3$ elle même conséquence de l'hypothèse $x > 0$.
Deuxième interprétation : On suppose $x > 0$. Alors l'implication "$y > 3 \Rightarrow z < 0$" est vraie.
Dans ce cas, on ne sait pas si $z < 0$ est vrai ou non. La seule chose qu'on sait c'est que sous l'hypothèse $x > 0$ et si on suppose $y > 3$ alors on aura $z < 0$. Mais rien ne dit que $y > 3$.

Exemple 2 (Exemple illustrant les points 5 et 6)**.** Soient X, Y, Z trois ensembles et soient $f : X \to Y$ et $g : Y \to Z$ deux applications.

1. Montrer que si $g \circ f$ est injective alors f est injective.
2. Montrer que si $g \circ f$ est surjective alors g est surjective.

Regardons la question 1. Je mets des commentaires en italique.
On suppose que $g \circ f$ est injective. *Ici le but est de montrer que f est injective autrement dit de montrer que pour tous $x, x' \in X$ tels que $f(x) = f(x')$, on a $x = x'$.*
Soient $x, x' \in X$ tels que $f(x) = f(x')$. *Là, on commence la démonstration du but*
En composant par g, on obtient $g(f(x)) = g(f(x'))$ c'est-à-dire $g \circ f(x) = g \circ f(x')$. Par hypothèse *(l'hypothèse intervient seulement maintenant)* $g \circ f$ est injective donc $x = x'$. *On a fini la démonstration du but.*

Voyons le point 2.
On suppose $g \circ f$ surjective.
Soit $z \in Z$ *(Là on démarre la démonstration du but, celui-ci consistant à montrer qu'il existe un antécédent $y \in Y$ à z.)*
Par hypothèse *(l'hypothèse intervient dès maintenant)* il existe $x \in X$ tel que $g \circ f(x) = z$. Posons $y = f(x)$. Par conséquent $y \in Y$ et $g(y) = z$.