UNIVERSITÉ PARIS 7
DENIS DIDEROT

MI3

Algèbre et analyse fondamentales I

CHAPITRE III

DÉTERMINANTS

année 2008-2009

Auteur : Thierry Joly

**Département de Formation
de 1er Cycle de Sciences Exactes**

DÉTERMINANTS

Plan du chapitre :

DÉTERMINANTS

1 Déterminants dans $\mathbb{R}^n$: une approche géométrique

1.1 Mesure des parallélotopes de $\mathbb{R}^n$ ($n \leqslant 3$)

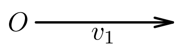
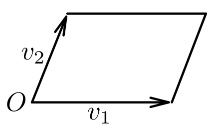
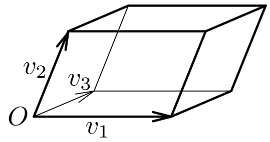
Pour n vecteurs quelconques $v_1, \ldots, v_n$ de l'espace $\mathbb{R}^n$ ($n \leqslant 3$), notons $P_{v_1,\ldots,v_n}$ l'ensemble

$$P_{v_1,\ldots,v_n} = \{M \; ; \; \exists x_1 \in [0,1] \ldots \exists x_n \in [0,1] \quad \overrightarrow{OM} = \sum_{i=1}^{n} x_i v_i\}.$$

Autrement dit, $P_{v_1,\ldots,v_n}$ est l'ensemble des points dont les coordonnées dans le repère $(O\,; v_1, \ldots, v_n)$ sont toutes comprises entre 0 et 1.

- Si les vecteurs $v_1, \ldots, v_n$ sont linéairement indépendants, autrement dit s'ils forment une *base* de $\mathbb{R}^n$, $P_{v_1,\ldots,v_n}$ est alors appelé *parallélotope engendré par les vecteurs* $v_1, \ldots, v_n$.
- Sinon (i.e. lorsque les vecteurs $v_1, \ldots, v_n$ sont liés), nous dirons que $P_{v_1,\ldots,v_n}$ est un *parallélotope dégénéré.*

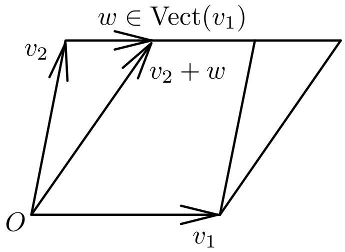
Parallélotopes

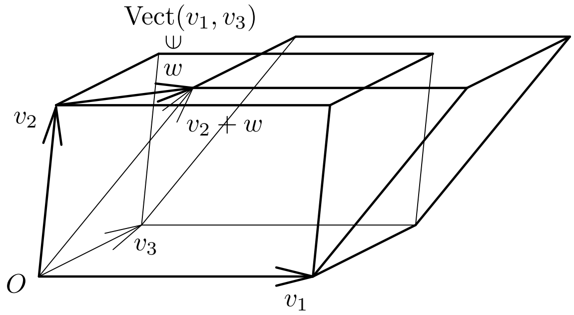
dim	nom	représentation	bord	mesure
1	segment		2 points	longueur (m)
2	parallélogramme		4 segments	aire (m^2)
3	parallélépipède		6 parallélogrammes	volume (m^3)

Notons $\mu(v_1, \ldots, v_n)$ la mesure du parallélotope $P_{v_1,\ldots,v_n}$ engendré par les vecteurs $v_1, \ldots, v_n$ (c'est-à-dire sa longueur, son aire ou son volume, selon la dimension n). Clairement, cette mesure est *nulle* si et seulement si $P_{v_1,\ldots,v_n}$ est *dégénéré* (i.e. s'il s'agit d'un segment réduit à un point, d'un parallélogramme "aplati" sur une droite ou encore d'un parallélépipède "aplati" dans un plan). Ainsi, $\mu(v_1, \ldots, v_n) = 0$ si et seulement si les vecteurs $v_1, \ldots, v_n$ sont liés.

Rassemblons d'autres propriétés intuitives de cette grandeur $\mu(v_1, \ldots, v_n)$:

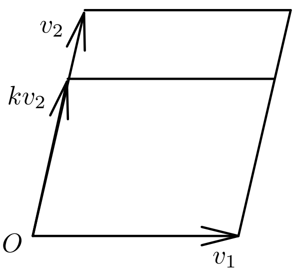
- Sa valeur est inchangée si l'on rajoute à l'un de ses arguments v_i un vecteur w appartenant au sous-espace engendré par ses autres arguments v_j $(j \neq i)$.

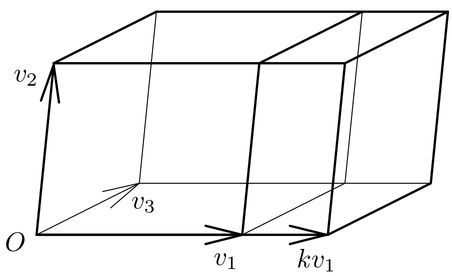




Autrement dit, la grandeur $\mu(v_1, \ldots, v_n)$ ne change pas si l'on rajoute à l'un des arguments v_i une *combinaison linéaire* $w = \sum_{j \neq i} k_j v_j$ quelconque des autres arguments v_j.

- Lorsque l'on multiplie l'un des arguments v_i par un réel *positif* k, la grandeur $\mu(v_1, \ldots, v_n)$ est elle-même multipliée par k.





En résumé, quelle que soit la dimension $n = 1, 2, 3$, on a les propriétés suivantes :

1. $\mu(v_1, \ldots, v_n) = 0$ si et seulement si les vecteurs $v_1, \ldots, v_n$ sont liés.
2. Pour toute combinaison linéaire $\sum_{j \neq i} k_j v_j$ des vecteurs $v_1, \ldots, v_n$ *autres que* v_i,
$$\mu(\ldots, v_i + \sum_{j \neq i} k_j v_j, \ldots) = \mu(\ldots, v_i, \ldots)$$
3. Pour tout argument v_i et tout réel $k > 0$,
$$\mu(\ldots, k\, v_i, \ldots) = k\, \mu(\ldots, v_i, \ldots)$$

Notons que cette dernière relation ne s'étend pas aux réels négatifs k, puisqu'alors son premier membre reste — par définition — positif, tandis que son second membre devient négatif.

1.2 Orientation des bases de $\mathbb{R}^n$ ($n \leqslant 3$) et déterminants

Un moyen simple d'étendre la propriété 3 ci-dessus aux réels négatifs k est d'affubler chaque mesure $\mu(v_1, \ldots, v_n) \neq 0$ d'un signe correspondant à l'*orientation de la base* $(v_1, \ldots, v_n)$. (Le système $(v_1, \ldots, v_n)$ est bien une *base* de $\mathbb{R}^n$ lorsque $\mu(v_1, \ldots, v_n) \neq 0$, en vertu de la propriété 1).

Intuitivement, deux bases $(v_1, \ldots, v_n)$ et $(v'_1, \ldots, v'_n)$ de $\mathbb{R}^n$ ont même *orientation* si l'on peut déformer continûment les vecteurs v_i jusqu'à les faire coïncider avec les vecteurs v'_i, de sorte que $(v_1, \ldots, v_n)$ reste une base de $\mathbb{R}^n$ tout au long de la déformation, autrement dit de telle sorte que le parallélotope $P_{v_1,\ldots,v_n}$ ne dégénère à aucun instant de la déformation.

Quelle que soit la dimension n, on a les faits suivants :

- Si $(v_1, \ldots, v_n)$ est une base quelconque de $\mathbb{R}^n$ et $(v'_1, \ldots, v'_n)$ la base obtenue en remplaçant *un seul* de ses vecteurs v_i par $-v_i$, alors les bases $(v_1, \ldots, v_n)$ et $(v'_1, \ldots, v'_n)$ n'ont pas la même orientation.
- Toute base de $\mathbb{R}^n$ a la même orientation que l'une de ces deux bases $(v_1, \ldots, v_n)$, $(v'_1, \ldots, v'_n)$.

Il en résulte que toutes les bases de $\mathbb{R}^n$ se répartissent selon exactement deux orientations. On a coutume d'appeler *directe* l'une des deux orientations possibles et *indirecte* l'autre. L'orientation directe est toujours définie par une règle *ad hoc* (par exemple "dans le sens de l'écriture" pour $\mathbb{R}$ vu comme une ligne horizontale, "dans le sens inverse des aiguilles d'une montre" pour $\mathbb{R}^2$, selon la règle "des trois doigts" ou encore "du bonhomme d'Ampère" pour $\mathbb{R}^3$). C'est un fait notable que les notions d'orientations directe et indirecte ne peuvent avoir de définition purement mathématique, *car il n'existe aucune propriété mathématique que possèderait une orientation et non l'autre.*

<u>*Déterminants dans* $\mathbb{R}^n$ ($n \leqslant 3$).</u> Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base quelconque de $\mathbb{R}^n$. On appelle *déterminant des vecteurs* $v_1, \ldots, v_n$ *par rapport à la base* $\mathcal{B}$, et l'on note $\det_{\mathcal{B}}(v_1, \ldots, v_n)$, le réel défini comme suit :

- Si les vecteurs $v_1, \ldots, v_n$ sont liés, alors : $\det_{\mathcal{B}}(v_1, \ldots, v_n) = \mu(v_1, \ldots, v_n) = 0$.
- Sinon, $(v_1, \ldots, v_n)$ est une base de l'espace et :
  - Le signe de $\det_{\mathcal{B}}(v_1, \ldots, v_n)$ est : $+$ si $(v_1, \ldots, v_n)$ a même orientation que la base $\mathcal{B}$,
$-$ sinon.
  - La valeur absolue du réel $\det_{\mathcal{B}}(v_1, \ldots, v_n)$ est la mesure $\mu(v_1, \ldots, v_n)$ du parallélotope $P_{v_1,\ldots,v_n}$ engendré par les vecteurs $v_1, \ldots, v_n$, en prenant comme unité la mesure du parallélotope $P_{e_1,\ldots,e_n}$:

$$\left|\det_{\mathcal{B}}(v_1, \ldots, v_n)\right| = \mu(v_1, \ldots, v_n) \neq 0.$$

Les propriétés des mesures $\mu(v_1, \ldots, v_n)$ mentionnées plus haut entraînent :

1. $\det_{\mathcal{B}}(v_1, \ldots, v_n) = 0$ si et seulement si les vecteurs $v_1, \ldots, v_n$ sont liés.
2. Pour toute combinaison linéaire $\sum_{j \neq i} k_j v_j$ des vecteurs $v_1, \ldots, v_n$ *autres que* v_i,

$$\det_{\mathcal{B}}(\ldots, v_i + \sum_{j \neq i} k_j v_j, \ldots) = \det_{\mathcal{B}}(\ldots, v_i, \ldots)$$

3. Pour tout argument v_i et tout réel k (*éventuellement négatif*),

$$\det_{\mathcal{B}}(\ldots, k\, v_i, \ldots) = k \det_{\mathcal{B}}(\ldots, v_i, \ldots)$$

La propriété 1 ci-dessus découle immédiatement de la propriété 1 des mesures $\mu(v_1, \ldots, v_n)$. Par ailleurs, une base $(v_1, \ldots, v_n)$ ne change pas d'orientation si l'on rajoute à l'un des vecteurs v_i une combinaison linéaire $w = \sum_{j \neq i} k_j v_j$ quelconque des autres vecteurs v_j. En effet, pour tout $t \in \mathbb{R}$, le système $(\ldots, v_i + tw, \ldots)$ est une base et l'application $t \mapsto (\ldots, v_i + tw, \ldots)$ décrit une transformation continue, au cours du temps t, permettant de passer de la base initiale $(\ldots, v_i, \ldots)$ (à l'instant $t = 0$) à la base $(\ldots, v_i + w, \ldots)$ (à l'instant $t = 1$). Ainsi, la propriété 2 ci-dessus se déduit de la propriété 2 des mesures $\mu(v_1, \ldots, v_n)$. Enfin, lorsque l'on remplace l'un des vecteurs v_i d'une base $(v_1, \ldots, v_n)$ par $k\, v_i$ ($k \neq 0$), cette base change d'orientation si et seulement si $k < 0$, de sorte que la propriété 3 ci-dessus se déduit à son tour de la propriété 3 des mesures $\mu(v_1, \ldots, v_n)$.

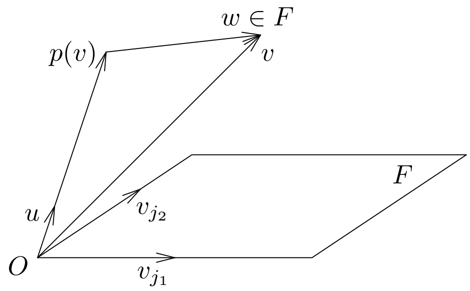
1.3 Définition formelle des déterminants dans $\mathbb{R}^n$

Commençons par dériver des propriétés 1, 2, 3 ci-dessus des déterminants d'autres propriétés algébriques plus célèbres.

En ne laissant varier qu'un seul argument $v_i \in \mathbb{R}^n$ de $\det_{\mathcal{B}}(v_1, \ldots, v_n)$ et en fixant tous les autres à des valeurs (vectorielles) quelconques, on définit une fonction partielle $\varphi : \mathbb{R}^n \to \mathbb{R}$. La propriété 3 des déterminants entraîne alors pour tout $v \in \mathbb{R}^n$ et tout réel k :

$$\varphi(k\, v) = k\, \varphi(v). \tag{1}$$

- Supposons que les $n-1$ vecteurs v_j ($j \neq i$) fixés dans la définition de φ sont linéairement indépendants. On peut alors choisir $u \in \mathbb{R}^n$ de façon à ce que les n vecteurs v_j ($j \neq i$) et u soient linéairement indépendants. Notons $p(v)$ la projection d'un vecteur $v \in \mathbb{R}^n$ quelconque sur la droite vectorielle engendré par u parallèlement au sous-espace F engendré les $n-1$ vecteurs fixés v_j :

  $(n = 3)$

  La propriété 2 des déterminants implique alors pour tout $v \in \mathbb{R}^n$:

  $$\varphi(p(v)) = \varphi(v). \tag{2}$$

  En effet, pour tout $v \in \mathbb{R}^n$, $p(v)$ s'obtient en rajoutant à v un vecteur w de F, c'est-à-dire une combinaison linéaire $w = \sum_{j \neq i} k_j v_j$ des vecteurs fixés v_j ($j \neq i$). Par ailleurs, pour tous vecteurs v, v' de $\mathbb{R}^n$, il existe des réels k, k' tels que :

  - $p(v) = ku$,
  - $p(v') = k'u$, et donc : (3)
  - $p(v + v') = (k + k')u$.

On déduit alors des relations (1), (2) et (3) :

$$\begin{aligned}\varphi(v+v') &\overset{(2)}{=} \varphi(p(v+v')) \overset{(3)}{=} \varphi((k+k')u) \overset{(1)}{=} (k+k')\varphi(u) = k\,\varphi(u) + k'\varphi(u)\\ &\overset{(1)}{=} \varphi(ku) + \varphi(k'u) \overset{(3)}{=} \varphi(p(v)) + \varphi(p(v')) \overset{(2)}{=} \varphi(v) + \varphi(v').\end{aligned}$$

- Supposons maintenant que les $n-1$ vecteurs v_j $(j \neq i)$ fixés dans la définition de φ sont liés. Pour tout $v \in \mathbb{R}^n$, les n vecteurs v_j $(j \neq i)$ et v sont alors évidemment liés et la propriété 1 des déterminants entraîne $\varphi(v) = 0$. Ainsi, on a $\varphi(v) = 0$ pour tout $v \in \mathbb{R}^n$, de sorte que la relation $\varphi(v+v') = \varphi(v) + \varphi(v')$ est encore vérifiée dans ce cas.

Ainsi, pour tous vecteurs v, v' de $\mathbb{R}^n$ et tout réel k :

- $\varphi(v+v') = \varphi(v) + \varphi(v')$,
- $\varphi(k\,v) = k\,\varphi(v)$.

Rappelons qu'une application φ vérifiant ces deux relations est dite *linéaire* (cf 1$^{\text{ère}}$ année) et que la *linéarité* d'une fonction φ s'exprime de façon équivalente par la seule relation :

$$\varphi(kv+v') = k\varphi(v) + \varphi(v'),$$

ou encore, pour toute combinaison linéaire $\displaystyle\sum_{j=1}^{m} k_j x_j$ de vecteurs x_j :

$$\varphi\Big(\sum_{j=1}^{m} k_j x_j\Big) = \sum_{j=1}^{m} k_j\,\varphi(x_j).$$

Toute application partielle de $\mathbb{R}^n$ dans $\mathbb{R}$ obtenue en fixant arbitrairement $n-1$ arguments de $\det_{\mathcal{B}}(v_1, \ldots, v_n)$ est donc linéaire :

$$\det_{\mathcal{B}}\Big(\ldots, \sum_{j=1}^{m} k_j x_j\,, \ldots\Big) = \sum_{j=1}^{m} k_j \det_{\mathcal{B}}(\ldots, x_j, \ldots).$$

On exprime ce fait en disant que l'application $\det_{\mathcal{B}}\ :\ \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ arguments}} \to \mathbb{R}$ est *linéaire en chacun de ses n arguments* ou encore qu'elle est *n-linéaire.*

Indiquons encore deux propriétés évidentes de cette application $\det_{\mathcal{B}}$. La première est un cas particulier de la propriété 1 vue plus haut : le déterminant s'annule chaque fois que deux de ses arguments sont égaux (car ses arguments sont alors évidemment liés).

$$\det_{\mathcal{B}}(\ldots, v, \ldots, v, \ldots) = 0$$

Une application vérifiant cette propriété est dite *alternée.* Enfin, de par le choix de l'unité de mesure fait, on a :

$$\det_{\mathcal{B}}(e_1, \ldots, e_n)\ =\ +\,\mu(P_{e_1,\ldots,e_n})\ =\ 1.$$

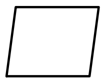
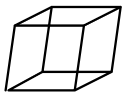
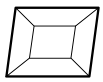
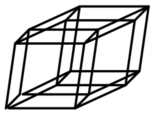
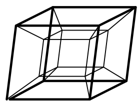
Dans la section suivante, nous démontrerons que ces trois dernières propriétés suffisent à caractériser l'application $\det_{\mathcal{B}}$. En d'autres termes, $\det_{\mathcal{B}}$ est *la seule* application $\underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ arguments}} \to \mathbb{R}$ qui soit n-linéaire, alternée et telle que $\det_{\mathcal{B}}(e_1, \ldots, e_n) = 1$.

Ce fait justifie le bien-fondé de la définition formelle suivante :

Définition Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base quelconque de $\mathbb{R}^n$. On appelle *déterminant par rapport à la base $\mathcal{B}$*, et l'on note $\det_{\mathcal{B}}$, l'application n-linéaire alternée $\det_{\mathcal{B}}\ :\ \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ arguments}} \to \mathbb{R}$ telle que $\det_{\mathcal{B}}(e_1, \ldots, e_n) = 1$.

Cette définition — comme toutes les définitions mathématiques — réduit à un ensemble minimal les propriétés de l'objet défini, au mépris de l'intuition que l'on s'en fait. Un intérêt majeur de cette pratique mathématique est de permettre la généralisation : remarquons que cette définition ne fait pas plus mention de la restriction $n \leqslant 3$ que la définition de *base de* $\mathbb{R}^n$ ou encore celle de *parallélotope*. On peut donc l'appliquer à toute dimension n. Elle permet alors de définir en retour pour tout n l'oriention d'une base de $\mathbb{R}^n$ ou la mesure d'un parallélotope de $\mathbb{R}^n$: l'orientation d'une base $(v_1, \ldots, v_n)$ de $\mathbb{R}^n$ par rapport à une autre base $\mathcal{B} = (e_1, \ldots, e_n)$ est le signe de $\det_{\mathcal{B}}(v_1, \ldots, v_n)$ et la mesure du parallélotope $P_{v_1,\ldots,v_n}$ est $\left|\det_{\mathcal{B}}(v_1, \ldots, v_n)\right|$.

Parallélotopes

dim	nom	représentation(s)	bord	mesure
0	point	•	—	1
1	segment	——	2 points	longueur (m)
2	parallélogramme		4 segments	aire (m^2)
3	parallélépipède	cavalière :  projective : 	6 parallélogrammes	volume (m^3)
4	parallélotope de dimension 4	 	8 parallélépipèdes	mesure (m^4)
$\vdots$				
n	parallélotope de dimension n		$2n$ parallélotopes de dimension $n-1$	mesure (m^n)
$\vdots$				

2 Déterminants dans un K-espace vectoriel ($\mathbf{K} = \mathbb{R}$ ou $\mathbb{C}$)

Nous allons maintenant étudier les déterminants dans le cadre plus général où l'ensemble des scalaires est indifféremment $\mathbb{R}$ ou $\mathbb{C}$. Dans toute la suite, $\mathbf{K}$ désignera l'un de ces deux ensembles.

2.1 Unicité et définition générale

Commençons par quelques propriétés des applications n-linéaires alternées :

Lemme 1 *Soit f une application n-linéaire alternée quelconque.*

(i). *La valeur prise par f reste inchangée lorsque l'on rajoute à l'un de ses arguments k fois un autre :*

$$f(v_1, \ldots, v_i, \ldots, v_n) = f(v_1, \ldots, v_i + kv_j, \ldots, v_n) \qquad (j \neq i).$$

(ii). *f change de signe lorsque l'on échange deux de ses arguments :*

$$f(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) = -f(v_1, \ldots, v_n).$$

Démonstration (i). En effet :

$$\begin{aligned} f(v_1,\ldots,v_i+kv_j,\ldots,v_j,\ldots,v_n) &= f(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_n)+k\,\underbrace{f(v_1,\ldots,v_j,\ldots,v_j,\ldots,v_n)}_{=0} \\ &= f(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_n). \end{aligned}$$

(ii). On applique (i) trois fois de suite :

$$\begin{aligned} f(\ldots,v_i,\ldots,v_j,\ldots) &\overset{1}{=} f(\ldots,v_i,\ldots,v_j-v_i,\ldots) \overset{2}{=} f(\ldots,v_i+(v_j-v_i),\ldots,v_j-v_i,\ldots) \\ &= f(\ldots,v_j,\ldots,v_j-v_i,\ldots) \overset{3}{=} f(\ldots,v_j,\ldots,-v_i,\ldots) = -f(\ldots,v_j,\ldots,v_i,\ldots). \end{aligned}$$

□

Théorème 2 (Unicité) *Soit E un $\mathbf{K}$-espace vectoriel de dimension n, $(e_1,\ldots,e_n)$ une base quelconque de cet espace E et f,g deux applications n-linéaires alternées de $\underbrace{E\times\cdots\times E}_{n \text{ arguments}}$ dans $\mathbb{R}$. Si $f(e_1,\ldots,e_n)=g(e_1,\ldots,e_n)$, alors $f=g$.*

Démonstration On établit la relation $f(v_1,\ldots,v_n)=g(v_1,\ldots,v_n)$ par récurrence sur le nombre p d'arguments v_i qui n'appartiennent pas à $\{e_1,\ldots,e_n\}$.

- Dans le cas $p=0$, tous les arguments v_i sont des vecteurs de la base $(e_1,\ldots,e_n)$. Si deux d'entre eux sont égaux, alors $f(v_1,\ldots,v_n)=0=g(v_1,\ldots,v_n)$. Sinon, $(v_1,\ldots,v_n)$ n'est autre que la base $(e_1,\ldots,e_n)$ à l'ordre des vecteurs près ; on peut alors ramener la suite d'arguments $(v_1,\ldots,v_n)$ à $(e_1,\ldots,e_n)$ par un certain nombre k d'échanges successifs d'arguments. En appliquant k fois de suite le lemme 1 (ii) à cette séquence d'échanges, on obtient : $f(v_1,\ldots,v_n)=-f(\ldots,v_j,\ldots,v_i,\ldots)=\ldots=(-1)^k f(e_1,\ldots,e_n)$ et de même : $g(v_1,\ldots,v_n)=(-1)^k g(e_1,\ldots,e_n)$, d'où $f(v_1,\ldots,v_n)=g(v_1,\ldots,v_n)$ puisque par hypothèse $f(e_1,\ldots,e_n)=g(e_1,\ldots,e_n)$.
- Supposons la relation établie lorsque p arguments ne sont pas des vecteurs de $\{e_1,\ldots,e_n\}$. Pour toute suite d'arguments $(v_1,\ldots,v_n)$ avec exactement $p+1$ vecteurs $v_i\notin\{e_1,\ldots,e_n\}$, soit $v_k=\sum_{i=1}^n x_i e_i$ l'un de ces $p+1$ vecteurs ; en écrivant $(v_1,\ldots,v_n)=(\ldots,v_k,\ldots)$, on a alors : $f(\ldots,v_k,\ldots)=f(\ldots,\sum_{i=1}^n x_i e_i,\ldots)=\sum_{i=1}^n f(\ldots,x_i e_i,\ldots)=\sum_{i=1}^n x_i f(\ldots,e_i,\ldots)$ et de même : $g(\ldots,v_k,\ldots)=\sum_{i=1}^n x_i g(\ldots,e_i,\ldots)$. Comme par hypothèse de récurrence, $f(\ldots,e_i,\ldots)=g(\ldots,e_i,\ldots)$ pour tout i, il s'ensuit : $f(\ldots,v_k,\ldots)=g(\ldots,v_k,\ldots)$.

□

En particulier, si $(e_1,\ldots,e_n)$ est une base d'un $\mathbf{K}$-espace vectoriel E et si f et g sont deux applications n-linéaires alternées de E^n dans $\mathbf{K}$ telles que $f(e_1,\ldots,e_n)=g(e_1,\ldots,e_n)=1$, alors $f=g$; autrement dit, il existe *au plus une* application n-linéaire alternée $f : E^n\to\mathbb{R}$ telle que $f(e_1,\ldots,e_n)=1$. Cela justifie la définition suivante :

Définition Soit E un $\mathbf{K}$-espace vectoriel de dimension n et $\mathcal{B}=(e_1,\ldots,e_n)$ une base quelconque de E. On appelle *déterminant par rapport à la base* $\mathcal{B}$, et l'on note $\det_{\mathcal{B}}$, l'application n-linéaire alternée $\det_{\mathcal{B}} : E^n\to\mathbf{K}$ telle que $\det_{\mathcal{B}}(e_1,\ldots,e_n)=1$.

2.2 Existence et calcul

Il reste à établir que cette application $\det_{\mathcal{B}}$ existe, ce qui peut sembler évident lorsque $\mathbf{K} = \mathbb{R}$ et $n \leqslant 3$, mais l'est beaucoup moins dans le cas général.

Théorème 3 (Existence) *Pour tout* $\mathbf{K}$*-espace vectoriel* E *et toute base* $\mathcal{B} = (e_1, \ldots, e_n)$ *de* E*, l'application* $\det_{\mathcal{B}} : E^n \to \mathbf{K}$ *existe bien. De plus, si* $n \geqslant 2$*,* $\mathcal{B}' = (e_1, \ldots, e_{i-1}, e_{i+1}, e_n)$ *et si* F *est le sous-espace de dimension* $n-1$ *dont une base est* $\mathcal{B}'$*, alors cette application se laisse définir à partir de l'application* $\det_{\mathcal{B}'} : F^{n-1} \to \mathbf{K}$ *par la relation :*

$$\det_{\mathcal{B}}(v_1, \ldots, v_n) = \sum_{j=1}^{n} (-1)^{i+j} f(v_j) \det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n)),$$

où $p : E \to F$ *désigne le projecteur sur* F *parallèlement au vecteur* e_i *et* $f : E \to \mathbf{K}$ *l'application qui à tout vecteur* $v = x_1 e_1 + \ldots + x_n e_n$ *de* E *associe sa* $i^{\text{ème}}$ *coordonnée* x_i *dans la base* $\mathcal{B}$*.*

Démonstration Démontrons l'existence de l'application $\det_{\mathcal{B}} : E^n \to \mathbf{K}$ par récurrence sur la dimension n de E.

Si $n = 1$, alors l'application $\varphi : E \to \mathbf{K}$ qui à tout vecteur $v = x_1 e_1$ de E associe son unique coordonnée x_1 est bien linéaire (donc 1-linéaire), trivialement alternée (puisque la situation où elle prend deux arguments égaux ne risque pas de se produire) et telle que $\varphi(e_1) = 1$, d'où $\varphi = \det_{\mathcal{B}}$ et l'existence de $\det_{\mathcal{B}}$.

Si $n \geqslant 2$, définissons $\mathcal{B}'$, F, p et f comme dans l'énoncé du théorème. Par hypothèse de récurrence, l'application déterminant $\det_{\mathcal{B}'} : F^{n-1} \to \mathbf{K}$ existe bien. Il nous reste à prouver que l'application $\varphi : E^n \to \mathbf{K}$ définie par :

$$\varphi(v_1, \ldots, v_n) = \sum_{j=1}^{n} (-1)^{i+j} f(v_j) \det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n))$$

est bien n-linéaire, alternée et telle que $\varphi(e_1, \ldots, e_n) = 1$; cela établira : $\varphi = \det_{\mathcal{B}}$ et donc l'existence de $\det_{\mathcal{B}}$.

• L'application $\varphi_j : E^n \to \mathbf{K}$ définie par :

$$\varphi_j(v_1, \ldots, v_n) = f(v_j) \det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n))$$

est linéaire en son argument v_j ; en effet, cela résulte de la linéarité de l'application f :

$$\begin{aligned}
\varphi_j\Big(\ldots, \underbrace{\sum_{r=1}^{m} k_r x_r}_{\uparrow\ j^{\text{ème}}\ \text{argument}}, \ldots\Big) &= f\Big(\sum_{r=1}^{m} k_r x_r\Big) \det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n)) \\
&= \Big(\sum_{r=1}^{m} k_r f(x_r)\Big) \det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n)) \\
&= \sum_{r=1}^{m} k_r f(x_r) \det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n)) \\
&= \sum_{r=1}^{m} k_r \varphi_j(\ldots, x_r, \ldots).
\end{aligned}$$

Par ailleurs, pour tout $j' \neq j$, la linéarité du projecteur p et la $(n-1)$-linéarité de $\det_{\mathcal{B}'}$ entraînent que φ_j est linéaire en son $j'^{\text{ème}}$ argument :

$$\begin{array}{rcl}
\varphi_j\Big(\underbrace{\ldots, \sum_{r=1}^{m} k_r x_r, \ldots}_{\substack{\uparrow \\ j'^{\text{ème}} \text{ argument}}}\Big) & = & f(v_j) \det_{\mathcal{B}'}\Big(\ldots, p\Big(\sum_{r=1}^{m} k_r x_r\Big), \ldots\Big) \\
& = & f(v_j) \det_{\mathcal{B}'}\Big(\ldots, \sum_{r=1}^{m} k_r p(x_r), \ldots\Big) \\
& = & f(v_j) \sum_{r=1}^{m} k_r \det_{\mathcal{B}'}(\ldots, p(x_r), \ldots) \\
& = & \sum_{r=1}^{m} k_r\, f(v_j) \det_{\mathcal{B}'}(\ldots, p(x_r), \ldots) \;=\; \sum_{r=1}^{m} k_r\, \varphi_j(\ldots, x_r, \ldots).
\end{array}$$

Ainsi, l'application φ_j est n-linéaire, et ce pour tout $j \in \{1, \ldots, n\}$. On en déduit facilement la n-linéarité de φ :

$$\begin{array}{rcl}
\varphi(\ldots, kv + v', \ldots) & = & \sum_{j=1}^{n} (-1)^{i+j}\, \varphi_j(\ldots, kv + v', \ldots) \\
& = & \sum_{j=1}^{n} (-1)^{i+j} \Big(k\, \varphi_j(\ldots, v, \ldots) + \varphi_j(\ldots, v', \ldots)\Big) \\
& = & k \sum_{j=1}^{n} (-1)^{i+j}\, \varphi_j(\ldots, v, \ldots) + \sum_{j=1}^{n} (-1)^{i+j}\, \varphi_j(\ldots, v', \ldots) \\
& = & k\, \varphi(\ldots, v, \ldots) + \varphi(\ldots, v', \ldots).
\end{array}$$

- Montrons à présent que l'application φ est alternée, i.e. :

$$r < r', \;\; v_r = v_{r'} \quad \Longrightarrow \quad \varphi(\ldots, v_r, \ldots, v_{r'}, \ldots) = 0,$$

par récurrence sur l'entier $r' - r$, autrement dit l'éloignement des arguments égaux v_r et $v_{r'}$.

  – Si $r' - r = 1$, alors ces arguments égaux sont les arguments successifs v_r et v_{r+1}. Dans ce cas, pour tout indice j autre que r et $r+1$, les vecteurs identiques $p(v_r)$ et $p(v_{r+1})$ apparaissent dans l'expression :

  $$\varphi_j(v_1, \ldots, v_n) = f(v_j) \det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n)) \;;$$

  comme l'application $\det_{\mathcal{B}'}$ est alternée, il s'ensuit $\varphi_j(v_1, \ldots, v_n) = 0$ pour tout $j \neq r, r+1$, et donc

  $$\begin{array}{rcl}
\varphi(v_1, \ldots, v_n) & = & (-1)^{i+r}\, \varphi_r(v_1, \ldots, v_n) + (-1)^{i+r+1}\, \varphi_{r+1}(v_1, \ldots, v_n) \\
& = & (-1)^{i+r}\big(\varphi_r(v_1, \ldots, v_n) - \varphi_{r+1}(v_1, \ldots, v_n)\big).
\end{array}$$

  De plus, l'égalité $v_r = v_{r+1}$ entraîne $f(v_r) = f(v_{r+1})$ et

  $$\det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{r-1}), p(v_{r+1}), \ldots, p(v_n)) = \det_{\mathcal{B}'}(p(v_1), \ldots, p(v_r), p(v_{r+2}), \ldots, p(v_n)),$$

  d'où : $\varphi_r(v_1, \ldots, v_n) = \varphi_{r+1}(v_1, \ldots, v_n)$, et par conséquent : $\varphi(v_1, \ldots, v_n) = 0$.

  – Si $r' - r \geqslant 2$, alors l'hypothèse de récurrence (pour l'éloignement $r' - r - 1$) entraîne :

  $$\varphi(\ldots, \underbrace{v_{r'-1}}_{\substack{\uparrow \\ r^{\text{ème}} \text{ argument}}}, \ldots, v_{r'-1}, v_{r'}, \ldots) = \varphi(\ldots, \underbrace{v_r + v_{r'-1}}_{\substack{\uparrow \\ r^{\text{ème}} \text{ argument}}}, \ldots, v_r + v_{r'-1}, v_{r'}, \ldots) = 0.$$

On en déduit à l'aide de la n-linéarité de φ (aux étapes $(*)$) :

$$
\begin{aligned}
&\varphi(\dots, v_r, \dots, v_{r'-1}, v_{r'}, \dots) \\
&\quad = \varphi(\dots, v_r, \dots, v_{r'-1}, v_{r'}, \dots) + \varphi(\dots, \overset{\substack{r^{\text{ème}}\text{ argument} \\ \downarrow}}{v_{r'-1}}, \dots, v_{r'-1}, v_{r'}, \dots) \\
&\quad \overset{(*)}{=} \varphi(\dots, v_r + v_{r'-1}, \dots, v_{r'-1}, v_{r'}, \dots) \\
&\quad \overset{(*)}{=} -\varphi(\dots, v_r + v_{r'-1}, \dots, v_r, v_{r'}, \dots) + \varphi(\dots, \underset{\substack{\uparrow \\ r^{\text{ème}}\text{ argument}}}{v_r + v_{r'-1}}, \dots, v_r + v_{r'-1}, v_{r'}, \dots) \\
&\quad = -\varphi(\dots, v_r + v_{r'-1}, \dots, v_r, v_{r'}, \dots).
\end{aligned}
$$

Enfin, l'égalité $v_r = v_{r'}$ et le cas $r' - r = 1$ traité plus haut permettent de conclure que le dernier membre est nul, d'où $\varphi(\dots, v_r, \dots, v_{r'}, \dots) = 0$.

• Établissons enfin la relation $\varphi(e_1, \dots, e_n) = 1$. Pour tout indice $j \neq i$, les définitions de l'application f et du projecteur p entraînent immédiatement $f(e_j) = 0$ et $p(e_j) = e_j$. Il s'ensuit :

$$
\begin{aligned}
\varphi(e_1, \dots, e_n) &= \sum_{j=1}^{n} (-1)^{i+j} f(e_j) \det_{\mathcal{B}'}(p(e_1), \dots, p(e_{j-1}), p(e_{j+1}), \dots, p(e_n)), \\
&= (-1)^{i+i} f(e_i) \det_{\mathcal{B}'}(p(e_1), \dots, p(e_{i-1}), p(e_{i+1}), \dots, p(e_n)), \\
&= f(e_i) \det_{\mathcal{B}'}(e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n).
\end{aligned}
$$

Comme par définition de f et de $\det_{\mathcal{B}'}$, on a : $f(e_i) = \det_{\mathcal{B}'}(e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n) = 1$, l'égalité $\varphi(e_1, \dots, e_n) = 1$ s'en déduit immédiatement.

$\square$

À partir de ces théorèmes d'existence et d'unicité, nous pouvons enfin établir en toute généralité la toute première propriété des déterminants :

Proposition 4 *Soit E un $\mathbf{K}$-espace vectoriel de dimension n muni d'une base $\mathcal{B}$. On a pour n vecteurs quelconques $v_1, \dots, v_n$ de E :*

$$v_1, \dots, v_n \text{ liés} \iff \det_{\mathcal{B}}(v_1, \dots, v_n) = 0.$$

Démonstration • Si $(v_1, \dots, v_n)$ est un système lié, alors il existe des scalaires non tous nuls k_i $(1 \leqslant i \leqslant n)$ tels que : $\sum_i k_i v_i = 0$ et l'un des vecteurs v_i est combinaison linéaire des autres. En effet, si mettons $k_i \neq 0$, alors : $k_i v_i = -\sum_{j \neq i} k_j v_j$ et en posant $k'_j = -k_j / k_i$ $(1 \leqslant j \leqslant n)$: $v_i = \sum_{j \neq i} k'_j v_j$. Il s'ensuit :

$$
\begin{aligned}
\det_{\mathcal{B}}(v_1, \dots, v_n) &= \det_{\mathcal{B}}\Big(v_1, \dots, v_{i-1}, \sum_{j \neq i} k'_j v_j, v_{i+1}, \dots, v_n\Big) \\
&= \sum_{j \neq i} k'_j \det_{\mathcal{B}}(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_n) = 0,
\end{aligned}
$$

puisque dans le déterminant de chaque terme de la dernière somme, le vecteur v_j est répété.
• Si en revanche $\mathcal{B}' = (v_1, \dots, v_n)$ est un système libre, c'est alors une base de E et le théorème d'existence nous permet de considérer l'application $\det_{\mathcal{B}'}$. L'application $f : E^n \to \mathbb{R}$ définie par :

$$f(x_1, \dots, x_n) = \det_{\mathcal{B}'}(e_1, \dots, e_n) \det_{\mathcal{B}}(x_1, \dots, x_n)$$

est dans ce cas une autre application n-linéaire alternée telle que $f(e_1, \dots, e_n) = \det_{\mathcal{B}'}(e_1, \dots, e_n)$, d'où, par le théorème d'unicité : $f = \det_{\mathcal{B}'}$. En faisant $(x_1, \dots, x_n) = (v_1, \dots, v_n)$ dans la définition de f, on obtient alors :

$$\det_{\mathcal{B}'}(e_1, \dots, e_n) \det_{\mathcal{B}}(v_1, \dots, v_n) = f(v_1, \dots, v_n) = \det_{\mathcal{B}'}(v_1, \dots, v_n) = 1,$$

ce qui implique $\det_{\mathcal{B}}(v_1, \dots, v_n) \neq 0$.

$\square$

Calcul des déterminants

La relation donnée dans l'énoncé du théorème d'existence fournit un procédé de calcul effectif des déterminants, puisqu'elle les ramène à des déterminants dans un espace de dimension inférieure : sachant calculer les déterminants dans un espace de dimension 1, on peut l'utiliser pour calculer les déterminants dans les espaces de dimension 2, puis l'utiliser encore pour le calcul des déterminants en dimension 3 et ainsi de suite.

Afin de calculer concrètement le scalaire $\det_{\mathcal{B}}(v_1, \ldots, v_n)$ pour des vecteurs donnés $v_1, \ldots, v_n$ d'un $\mathbf{K}$-espace vectoriel de base $\mathcal{B} = (e_1, \ldots, e_n)$, on écrit en colonnes les coordonnées de chacun de ces vecteurs dans la base $\mathcal{B}$:

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \quad \begin{pmatrix} a_{12} \\ \vdots \\ a_{n2} \end{pmatrix}, \quad \ldots \quad \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix},$$

où $v_1 = \sum_{i=1}^{n} a_{i1} e_i,\ v_2 = \sum_{i=1}^{n} a_{i2} e_i, \ldots, v_n = \sum_{i=1}^{n} a_{in} e_i$, puis on rassemble ces colonnes entre deux traits verticaux pour représenter le déterminant à calculer :

$$\det_{\mathcal{B}}(v_1, \ldots, v_n) = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

On l'appelle alors aussi *déterminant de la matrice carrée* $n \times n$:

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

et on le note $\det A$.

Transcrivons maintenant la relation du théorème d'existence dans cette notation matricielle : la projection de $v_r = \sum_q a_{q,r}\, e_q$ sur le sous-espace F engendré par $\mathcal{B}' = (e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n)$ parallèlement au vecteur e_i est le vecteur $p(v_r) = \sum_{q \neq i} a_{q,r}\, e_q$, dont les coordonnées dans la base $\mathcal{B}'$ s'écrivent :

$$\begin{pmatrix} a_{1,r} \\ \vdots \\ a_{i-1,r} \\ a_{i+1,r} \\ \vdots \\ a_{n,r} \end{pmatrix}.$$

On en déduit :

$$\det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n)) = \begin{vmatrix} a_{1,1} & \ldots & a_{1,j-1} & a_{1,j+1} & \ldots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \ldots & a_{n,j-1} & a_{1,j+1} & \ldots & a_{1,n} \end{vmatrix}.$$

Ainsi, $\det(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n))$ est le déterminant de la matrice obtenue en rayant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne de la matrice A :

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,j-1} & a_{i,j} & a_{i,j+1} & \cdots & a_{i,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j} & a_{1,j+1} & \cdots & a_{1,n} \end{pmatrix}$$

Comme cette manipulation sera fréquente, on adopte dans toute la suite la notation suivante :

Notation Pour toute matrice carrée $n \times n$ A et tous $i, j \in \{1, \ldots, n\}$, on note A_{ij} la matrice carrée $(n-1) \times (n-1)$ obtenue à partir de A en supprimant sa $i^{\text{ème}}$ ligne et sa $j^{\text{ème}}$ colonne.

On a donc : $\det(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n)) = \det A_{ij}$. Par ailleurs, la $i^{\text{ème}}$ coordonnée de v_j dans la base $\mathcal{B} = (e_1, \ldots, e_n)$ est $f(v_j) = a_{ij}$, de sorte que la relation :

$$\det_{\mathcal{B}}(v_1, \ldots, v_n) = \sum_{j=1}^{n} (-1)^{i+j} f(v_j) \det_{\mathcal{B}'}(p(v_1), \ldots, p(v_{j-1}), p(v_{j+1}), \ldots, p(v_n)),$$

se traduit par :

$$\boxed{\det A = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij} \det A_{ij}.}$$

Cette nouvelle forme de la relation est appelée *formule de développement selon la $i^{\text{ème}}$ ligne*.

À des fins d'utilisation pratique, on peut l'exprimer plus naïvement par :

$$\det A = \pm\, a_{i1} \det A_{i1} \pm a_{i2} \det A_{i2} \pm \ldots \pm a_{in} \det A_{in},$$

où les signes placés devant les coefficients a_{ij} sont déterminés comme sur un damier par leur position dans la matrice :

$$A = \begin{pmatrix} + & - & + & - & \cdots\cdots \\ - & + & - & + & \cdots\cdots \\ + & - & + & - & \cdots\cdots \\ - & + & - & + & \cdots\cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

(Attention : commencer par un signe + en haut à gauche ou en bas à droite.)

L'importance de cette formule motive la terminologie suivante :

Définition Soit $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ une matrice carrée quelconque.

- On appelle *mineur associé au coefficient* a_{ij} le déterminant $\det A_{ij}$.
- On appelle *cofacteur du coefficient* a_{ij} le scalaire $(-1)^{i+j} \det A_{ij}$.
- On appelle *comatrice de A* et l'on note $\operatorname{Com} A$ la matrice $(c_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$, où c_{ij} est le cofacteur du coefficient a_{ij} de A.

Déterminants d'ordre 1.

Comme remarqué dans la preuve du théorème d'existence, si E est un espace de dimension 1, alors le déterminant $\det_{\mathcal{B}}(v)$ d'un vecteur $v \in E$ par rapport à une base $\mathcal{B} = (e_1)$ de E est son unique coordonnée dans cette base. Il s'ensuit que le déterminant d'une matrice 1×1 :

$$A = \Big(a_{11} \Big)$$

est son unique coefficient : $\det A = a_{11}$. (La notation $\Big| a_{11} \Big|$ de ce déterminant est fortement déconseillée pour une raison évidente.)

Déterminants d'ordre 2.

En développant selon la première ligne le déterminant d'une matrice :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

on obtient : $\det A = \ + a_{11} \det A_{11} - a_{12} \det A_{12} = a_{11}a_{22} - a_{12}a_{21}$. On retrouve ainsi la fameuse règle "du produit en croix". (En développant $\det A$ selon la seconde ligne, on obtient naturellement le même résultat : $\det A = \ - a_{21} \det A_{21} + a_{22} \det A_{22} = -a_{21}a_{12} + a_{22}a_{11}$.)

Déterminants d'ordre 3.

Le développement selon la première ligne du déterminant d'une matrice :

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

donne :

$$\begin{aligned} \det A \ &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}\Big(a_{22}a_{33} - a_{23}a_{32}\Big) - a_{12}\Big(a_{21}a_{33} - a_{23}a_{31}\Big) + a_{13}\Big(a_{21}a_{32} - a_{22}a_{31}\Big) \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned}$$

Les 6 termes de cette dernière expression peuvent être retrouvés en répétant les deux premières colonnes de la matrice A à sa suite, puis en regroupant les coefficients selon des diagonales de la façon suivante :

$$\begin{array}{ccccc} + & + & + & & \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ - & - & - & & \end{array}$$

Cette règle, dite *"de Sarrus"*, semble généraliser la règle du produit en croix des déterminants d'ordre 2, mais n'a pas d'équivalent aux ordres supérieurs. Par ailleurs, elle est inutilement coûteuse en calculs, puisqu'elle nécessite 12 multiplications et 5 additions, tandis que le calcul direct du développement selon une ligne ne requiert que 9 multiplications et 5 additions (cf plus haut).

La règle de Sarrus est un bel exemple de règle esthétique mais parfaitement inutile d'un point de vue calculatoire. La théorie des déterminants en fourmille. Indiquons un autre exemple célèbre d'élégance théorique menant à des calculs catastrophiques : la méthode de résolution des systèmes d'équations linéaires appelée *Méthode de Cramer*.

Pour tout système d'équations

$$\left[\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & \cdots & + & a_{nn}x_n & = & b_n \end{array}\right. \qquad (S)$$

le déterminant

$$\Delta_S = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

est appelé *déterminant du système* (S) et l'on nomme *déterminant* Δ_{x_i} *associé à l'inconnue* x_i le déterminant obtenu en remplaçant la colonne de Δ_S correspondant à x_i (autrement dit sa $i^{\text{ème}}$ colonne) par la colonne des seconds membres :

$$\Delta_{x_i} = \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}$$
$$\uparrow$$
$$i^{\text{ème}} \text{ colonne}$$

Règle de Cramer Si $\Delta_S \neq 0$, alors le système (S) possède une seule solution $(x_1, \ldots, x_n)$ et cette unique solution est donnée par les relations :

$$x_i = \frac{\Delta_{x_i}}{\Delta_S} \qquad (1 \leqslant i \leqslant n).$$

Afin de justifier cette règle, considérons les vecteurs de $\mathbf{K}^n$:

$$v_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix} \quad (1 \leqslant i \leqslant n) \qquad \text{et} \qquad w = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Le système (S) s'exprime alors sous la forme plus synthétique :

$$x_1v_1 \ + \ \cdots \ + \ x_nv_n \ = \ w. \qquad (S)$$

De plus, on a : $\Delta_S = \det_{\mathcal{B}}(v_1, \ldots, v_n)$, où $\mathcal{B}$ désigne la base canonique de l'espace $\mathbf{K}^n$. En vertu de la proposition 4, l'hypothèse $\Delta_S \neq 0$ entraîne donc que le système $(v_1, \ldots, v_n)$ est une base de $\mathbf{K}^n$; il s'ensuit que w — comme tout autre vecteur de $\mathbf{K}^n$ — peut s'écrire comme une combinaison linéaire des vecteurs $v_1, \ldots, v_n$, autrement dit l'équation (S) possède une solution $(x_1, \ldots, x_n)$. Enfin, une solution quelconque $(x_1, \ldots, x_n)$ de (S) vérifie :

$$\begin{aligned} \Delta_{x_i} &= \det_{\mathcal{B}}(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n) \\ &= \det_{\mathcal{B}}\Big(v_1, \ldots, v_{i-1}, \sum_{j=1}^{n} x_jv_j, v_{i+1}, \ldots, v_n\Big) \\ &= \sum_{j=1}^{n} x_j \det_{\mathcal{B}}(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_n). \end{aligned}$$

Chaque terme de cette dernière somme tel que $j \neq i$ est nul, car le vecteur v_j est répété dans son déterminant. On a donc : $\Delta_{x_i} = x_i \det_{\mathcal{B}}(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) = x_i\Delta_S$, d'où $x_i = \Delta_{x_i}/\Delta_S$ pour tout i et l'unicité de la solution.

La méthode de Cramer a été exposée ici à titre d'exercice théorique, bien sûr, mais aussi en forme de mise en garde : *À cause de la lourdeur des calculs de déterminants, les relations qui les invoquent sont généralement inefficaces pour mener concrètement des calculs* ; l'intérêt de ces relations est le plus souvent d'ordre théorique. Ainsi, la méthode de Cramer établit rapidement l'existence et l'unicité des systèmes linéaires $n \times n$ de déterminant non nul ; en revanche, sa mise en œuvre nécessite le calcul de $n+1$ déterminants d'ordre n, c'est-à-dire bien plus de calculs que par exemple la méthode du pivot de Gauss.

Quant aux calculs de déterminants proprement dits (qui s'avèrent parfois inévitables), le jeu de règles données plus loin (cf 2.4 Opérations sur les déterminants) permettra de les effectuer à moindre frais.

2.3 Déterminants et transposition

Définition On appelle *transposée d'une matrice* $n \times p$:

$$A = \begin{pmatrix} a_{11} \ldots\ldots a_{1p} \\ \vdots \qquad\quad \vdots \\ a_{n1} \ldots\ldots a_{np} \end{pmatrix} \qquad \text{la matrice } p \times n : \qquad {}^tA = \begin{pmatrix} a_{11} \ldots a_{n1} \\ \vdots \qquad \vdots \\ \vdots \qquad \vdots \\ a_{p1} \ldots a_{np} \end{pmatrix}$$

obtenue à partir de A en interchangeant lignes et colonnes.

Proposition 5 *Pour toutes matrices* A, B *:*

(i). ${}^t({}^tA) = A$.

(ii). ${}^t(AB) = {}^tB\,{}^tA$,

Démonstration (i) est évident. (ii) résulte de la définition du produit matriciel : le produit AB d'une matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}$ par une matrice $B = (b_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}}$ est la matrice $C = (c_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant q}}$ dont les coefficients sont :

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

Le produit ${}^tB\,{}^tA$ de la matrice ${}^tB = (b'_{ij})_{\substack{1\leqslant i\leqslant q\\1\leqslant j\leqslant p}}$ par la matrice ${}^tA = (a'_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ est alors de même la matrice $C' = (c'_{ij})_{\substack{1\leqslant i\leqslant q\\1\leqslant j\leqslant n}}$ dont les coefficients sont :

$$c'_{ij} = \sum_{k=1}^{p} b'_{ik} a'_{kj} = \sum_{k=1}^{p} b_{ki} a_{jk} = c_{ji},$$

autrement dit $C' = {}^tC$, soit encore ${}^tB\,{}^tA = {}^t(AB)$.

□

Afin d'établir le lemme suivant, commençons par remarquer qu'une matrice carrée possédant deux lignes identiques a un déterminant nul. En effet, les colonnes d'une telle matrice peuvent être vues comme n vecteurs $v_1, \ldots, v_n$ de l'espace $\mathbf{K}^n$ et son déterminant comme le déterminant $\det_{\mathcal{B}}(v_1, \ldots, v_n)$ par rapport à la base canonique $\mathcal{B} = (e_1, \ldots, e_n)$ de $\mathbf{K}^n$. Dire que les $i^{\text{ème}}$ et $i'^{\text{ème}}$ lignes de la matrice sont identiques ($i \neq i'$) revient à dire que les $i^{\text{ème}}$ et $i'^{\text{ème}}$ coordonnées dans $\mathcal{B}$ de chaque vecteur v_j sont égales, autrement dit que chaque vecteur v_j appartient au sous-espace F engendré par $e_i + e_{i'}$ et les $n-2$ vecteurs e_k autres que $e_i, e_{i'}$. Comme $\dim F = n-1$, les n vecteurs v_j sont donc liés, d'où par la proposition 4 : $\det_{\mathcal{B}}(v_1, \ldots, v_n) = 0$.

Lemme 6 *Pour tout matrice $n \times n$ A, on a :*

$$A\,{}^t\mathrm{Com}\,A \;=\; \det A \,.\, I_n,$$

où I_n désigne la matrice identité $n \times n$.

Démonstration Posons $\mathrm{Com}\,A = C = (c_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ et $A\,{}^t\mathrm{Com}\,A = A\,{}^tC = (d_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$. Par définition de la comatrice, on a : $c_{ij} = (-1)^{i+j}\det A_{ij}$, d'où pour tous $k, i \in \{1, \ldots, n\}$:

$$d_{ki} = \sum_{j=1}^{n} a_{kj}c_{ij} = \sum_{j=1}^{n}(-1)^{i+j}\, a_{kj}\det A_{ij}.$$

- Si $k = i$, on reconnaît la formule du développement selon la $i^{\text{ème}}$ ligne du déterminant de A, d'où : $d_{ii} = \det A$.
- Si $k \neq i$, il s'agit alors du développement selon la $i^{\text{ème}}$ ligne du déterminant de la matrice $A' = (a'_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ obtenue en remplaçant la $i^{\text{ème}}$ ligne de A par sa $k^{\text{ème}}$ ligne. En effet, on a bien $a'_{ij} = a_{kj}$ et $A'_{ij} = A_{ij}$ pour tout j. Comme la matrice A' possède deux lignes identiques (la $i^{\text{ème}}$ et la $k^{\text{ème}}$), la remarque ci-dessus permet d'en déduire : $d_{ki} = \det A' = 0$.

Ainsi, la matrice $A\,{}^t\mathrm{Com}\,A = (d_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ est bien égale à la matrice $\det A \,.\, I_n$.

□

Proposition 7 *Pour toute matrice $n \times n$ A, on a :* $\det {}^tA = \det A$.

Démonstration Par récurrence sur l'entier n. Pour $n = 1$, c'est évident puisqu'alors ${}^tA = A$. Supposons donc la proposition établie au rang $n-1$ et montrons-la au rang n. Par définition, le coefficient à l'intersection de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne de ${}^t\mathrm{Com}\,A$ est $c_{ij} = (-1)^{j+i}\det A_{ji}$. En remarquant que la matrice $({}^tA)_{ij}$ obtenue en supprimant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne de ${}^tA$ est exactement la matrice $(n-1)\times(n-1)$ ${}^tA_{ji}$, on obtient par hypothèse de récurrence $\det ({}^tA)_{ij} = \det {}^tA_{ji} = \det A_{ji}$, d'où : $c_{ij} = (-1)^{i+j}\det ({}^tA)_{ij}$, ce qui établit :

$${}^t\mathrm{Com}\,A = \mathrm{Com}\,{}^tA. \tag{$*$}$$

Par ailleurs, comme la matrice $\det {}^tA \,.\, I_n$ est naturellement égale à sa transposée, on a aussi $A\,{}^t(\det {}^tA \,.\, I_n) \;=\; A\,(\det {}^tA \,.\, I_n) \;=\; \det {}^tA \,.\, A\,I_n \;=\; \det {}^tA \,.\, A$. On en déduit à l'aide de la proposition 5 et du lemme 6 :

$$\begin{aligned}\det {}^tA \,.A \;&=\; A\,{}^t(\det {}^tA \,.\, I_n) \;\overset{6}{=}\; A\,{}^t({}^tA\,{}^t\mathrm{Com}\,{}^tA) \;\overset{5(\mathrm{ii})}{=}\; A\,{}^t({}^t\mathrm{Com}\,{}^tA)\,{}^t({}^tA)\\ &\overset{5(\mathrm{i})}{=}\; A\,\mathrm{Com}\,{}^tA\;A \;\overset{(*)}{=}\; (A\,{}^t\mathrm{Com}\,A)\,A \;\overset{6}{=}\; (\det A \,.\, I_n)A \;=\; \det A \,.\, A.\end{aligned}$$

Cette égalité entraîne évidemment pour chaque coefficient a_{ij} de la matrice A :

$$\det {}^tA \,.\, a_{ij} = \det A \,.\, a_{ij}.$$

- Si l'un de ces coefficients a_{ij} est non nul, il s'ensuit : $\det {}^tA \;=\; \det A$.
- Sinon, on a ${}^tA \;=\; A \;=\; 0$, d'où là encore : $\det {}^tA \;=\; \det A$.

□

2.4 Opérations sur les déterminants

On se contente ici de rassembler toutes les règles de manipulation des déterminants.

Proposition 8 *Soit* $A = \begin{pmatrix} a_{11} \dots a_{1n} \\ \vdots \qquad \vdots \\ a_{n1} \dots a_{nn} \end{pmatrix}$ *une matrice* $n \times n$ *quelconque.*

(i). *Si* $n = 1$, *alors* $\det A$ *est l'unique coefficient de* A *:* $\det A = a_{11}$.

(ii). *Si* $n \geqslant 2$, *alors pour tous* $i, j \in \{1, \dots, n\}$ *:*

- $\det A = \displaystyle\sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$ *(développement selon la* $i^{\text{ème}}$ *ligne)*
- $\det A = \displaystyle\sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$ *(développement selon la* $j^{\text{ ème}}$ *colonne).*

Autrement dit :

$$\begin{aligned} \det A &= \pm a_{i1} \det A_{i1} \pm a_{i2} \det A_{i2} \pm \dots \pm a_{in} \det A_{in}, \\ \det A &= \pm a_{1j} \det A_{1j} \pm a_{2j} \det A_{2j} \pm \dots \pm a_{nj} \det A_{nj}, \end{aligned}$$

où les signes alternés placés devant les coefficients a_{ij} *sont déterminés par la position de ces derniers dans la matrice à l'aide du diagramme :*

$$A = \begin{pmatrix} + & - & + & - & \cdots\cdots \\ - & + & - & + & \cdots\cdots \\ + & - & + & - & \cdots\cdots \\ - & + & - & + & \cdots\cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

(Attention : commencer par un signe + en haut à gauche ou en bas à droite.)

(iii). *On factorise* $\det A$ *par un scalaire* k *en factorisant par* k *une* seule *ligne ou une* seule *colonne de* A *:*

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = k \begin{pmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{pmatrix} \quad \Longrightarrow \quad \det A = k \begin{vmatrix} a_{11} \dots x_{1j} \dots a_{1n} \\ \vdots \qquad \vdots \qquad \vdots \\ a_{n1} \dots x_{nj} \dots a_{nn} \end{vmatrix}$$

$$\begin{pmatrix} a_{i1} & \dots\dots & a_{in} \end{pmatrix} = k \begin{pmatrix} y_{i1} & \dots\dots & y_{in} \end{pmatrix} \quad \Longrightarrow \quad \det A = k \begin{vmatrix} a_{11} \dots\dots\dots a_{1n} \\ \vdots \qquad\qquad \vdots \\ y_{11} \dots\dots\dots y_{1n} \\ \vdots \qquad\qquad \vdots \\ a_{n1} \dots\dots\dots a_{nn} \end{vmatrix}$$

(iv). $\det A$ *ne change pas si l'on rajoute à une colonne de* A k *fois une autre colonne de* A *ou si l'on rajoute à une ligne* k *fois une autre ligne :*

$$\det A = \begin{vmatrix} a_{11} \dots\dots a_{1j} + k a_{1j'} \dots\dots a_{1n} \\ a_{21} \dots\dots a_{2j} + k a_{2j'} \dots\dots a_{2n} \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{n1} \dots\dots a_{nj} + k a_{nj'} \dots\dots a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} \quad \dots\dots\dots\dots \quad a_{1n} \\ \vdots \qquad\qquad\qquad \vdots \\ a_{i1} + k a_{i'1} \quad \dots\dots \quad a_{in} + k a_{i'n} \\ \vdots \qquad\qquad\qquad \vdots \\ a_{n1} \quad \dots\dots\dots\dots \quad a_{nn} \end{vmatrix}$$

(v). $\det A$ *change de signe si l'on échange deux colonnes ou deux lignes de* A.

(vi). $\det A = \det {}^t\!A$.

Démonstration On a déjà démontré la plupart de ces règles. (i) est un rappel de la partie *Calcul des déterminants* (cf 2.2 Existence et calcul). (vi) est exactement la proposition 7. Chacune des autres règles (ii)-(v) consiste en une paire d'énoncés, l'un s'obtenant à partir de l'autre en intervertissant les rôles des lignes et des colonnes. Chaque fois, les deux énoncés sont équivalents en vertu de (vi), de sorte qu'il suffit d'établir l'un d'entre eux. En effet, les lignes de la matrice A correspondent aux colonnes de ${}^tA$ et réciproquement. Ainsi, la formule de développement selon la $j^{\text{ème}}$ colonne de la règle (ii) se déduit facilement de la formule de développement selon une ligne établie plus haut dans la partie *Calcul des déterminants* : En posant ${}^tA = A' = (a'_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$, le développement selon la $j^{\,ème}$ *ligne* de $\det A'$ s'écrit :

$$\det A' = \sum_{i=1}^{n} (-1)^{j+i}\, a'_{ji} \det A'_{ji}.$$

En remarquant que l'on a $a'_{ji} = a_{ij}$ et $A'_{ji} = {}^tA_{ij}$ pour tout indice i, (vi) permet de conclure :

$$\det A = \det A' = \sum_{i=1}^{n} (-1)^{j+i}\, a'_{ji} \det A'_{ji} = \sum_{i=1}^{n} (-1)^{i+j}\, a_{ij} \det A_{ij}.$$

Enfin, les règles (iii), (iv) et (v) restreintes aux seules manipulations de colonnes reformulent dans la notation matricielle des propriétés générales des applications n-linéaires alternées : (iii) exprime partiellement la n-linéarité des déterminants, (iv) et (v) sont une reformulation du lemme 1.

□

Commentaire sur l'utilité de ces règles

Dans le calcul pratique de déterminants, on utilise essentiellement (i) et (ii). On se sert aussi abondamment de (iv) pour écourter les calculs en faisant apparaître des 0 sur la ligne ou la colonne selon laquelle on développe le déterminant. (iv) permet notamment d'utiliser systématiquement la méthode du pivot de Gauss pour ne laisser qu'un coefficient non nul dans la ligne ou colonne de développement. (iii) est parfois utilisé à l'occasion d'une factorisation évidente de colonne ou de ligne dans le but de simplifier les coefficients. Les propriétés (v) et (vi) sont délaissées, car elles ne permettent pas de faire progresser le calcul.

Exemple de calcul

$$\begin{aligned}
\begin{vmatrix} -1 & 1 & -3 & 2 \\ 3 & 2 & 11 & 5 \\ 7 & -4 & 2 & -8 \\ 0 & 1 & -1 & 2 \end{vmatrix}
&= \begin{vmatrix} -1 & 1 & -3 & 0 \\ 3 & 2 & 11 & 1 \\ 7 & -4 & 2 & 0 \\ 0 & 1 & -1 & 0 \end{vmatrix} && \text{en ajoutant } -2 \text{ fois la } 2^{\text{ème}} \text{ col. à la } 4^{\text{ème}} \text{ (iv),} \\
&= \begin{vmatrix} -1 & 1 & -3 \\ 7 & -4 & 2 \\ 0 & 1 & -1 \end{vmatrix} && \text{en développant selon la } 4^{\text{ème}} \text{ colonne (ii),} \\
&= \begin{vmatrix} -1 & 1 & -2 \\ 7 & -4 & -2 \\ 0 & 1 & 0 \end{vmatrix} && \text{en ajoutant la } 2^{\text{ème}} \text{ colonne à la } 3^{\text{ème}} \text{ (iv),} \\
&= -\begin{vmatrix} -1 & -2 \\ 7 & -2 \end{vmatrix} && \text{en développant selon la } 3^{\text{ème}} \text{ ligne (ii),} \\
&= 2\begin{vmatrix} -1 & 1 \\ 7 & 1 \end{vmatrix} && \text{en factorisant la } 2^{\text{ème}} \text{ colonne (ii),} \\
&= 2\,(-1-7) = -16 && \text{en développant selon la } 1^{\text{ère}} \text{ colonne (ii) ou} \\
& && \text{en utilisant la formule des déterminants } 2\times 2.
\end{aligned}$$

3 Déterminant d'un endomorphisme

3.1 Définition et multiplicativité

Commençons par une conséquence importante du théorème d'unicité (théorème 2) : Soit E un $\mathbf{K}$-espace vectoriel de dimension n et $\mathcal{B} = (e_1, \ldots, e_n)$ une base quelconque de E. *Alors pour toute application n-linéaire alternée $\varphi : E^n \to \mathbf{K}$, il existe une constante $k \in \mathbf{K}$ telle que φ soit de la forme :*

$$\varphi(v_1, \ldots, v_n) = k \det_{\mathcal{B}}(v_1, \ldots, v_n).$$

En effet, pour tout scalaire k, on vérifie facilement que l'application $\varphi' : E^n \to \mathbf{K}$ définie par : $\varphi'(v_1, \ldots, v_n) = k \det_{\mathcal{B}}(v_1, \ldots, v_n)$ est n-linéaire alternée. En choisissant k égal à $\varphi(e_1, \ldots, e_n)$, on a alors de plus : $\varphi(e_1, \ldots, e_n) = \varphi'(e_1, \ldots, e_n)$, d'où par le théorème d'unicité : $\varphi = \varphi'$.

Soit maintenant f un endomorphisme de E quelconque (c'est-à-dire une application linéaire $f : E \to E$). L'application $\varphi : E^n \to \mathbf{K}$ définie par :

$$\varphi(v_1, \ldots, v_n) = \det_{\mathcal{B}}(f(v_1), \ldots, f(v_n))$$

est alors n-linéaire alternée. Cela résulte de la linéarité de f et du fait que $\det_{\mathcal{B}}$ est elle-même n-linéaire alternée :

$$\begin{aligned}
\varphi(\ldots, kv_i + v'_i, \ldots) &= \det_{\mathcal{B}}(\ldots, f(kv_i + v'_i), \ldots)\\
&= \det_{\mathcal{B}}(\ldots, k\, f(v_i) + f(v'_i), \ldots)\\
&= k \det_{\mathcal{B}}(\ldots, f(v_i), \ldots) + \det_{\mathcal{B}}(\ldots, f(v'_i), \ldots)\\
&= k\, \varphi(\ldots, v_i, \ldots) + \varphi(\ldots, v'_i, \ldots),
\end{aligned}$$

$$i \neq i',\ v_i = v_{i'} \quad \Longrightarrow \quad \varphi(\ldots, v_i, \ldots, v_{i'}, \ldots) = \det_{\mathcal{B}}(\ldots, f(v_i), \ldots, f(v_i), \ldots) = 0.$$

En vertu de la conséquence du théorème d'unicité tirée ci-dessus, il s'ensuit que φ est de la forme : $\varphi(v_1, \ldots, v_n) = k \det_{\mathcal{B}}(v_1, \ldots, v_n)$ pour une certaine constante $k \in \mathbf{K}$; autrement dit, pour tous vecteurs $v_1, \ldots, v_n$ de E :

$$\det_{\mathcal{B}}(f(v_1), \ldots, f(v_n)) = k \det_{\mathcal{B}}(v_1, \ldots, v_n). \tag{$*$}$$

De plus, cette constante k ne dépend même pas de la base $\mathcal{B}$ choisie ! En effet, toujours en vertu de la même remarque faite ci-dessus, pour toute autre base $\mathcal{B}'$, l'application n-linéaire alternée $\det_{\mathcal{B}'}$ est nécessairement de la forme : $\det_{\mathcal{B}'}(v_1, \ldots, v_n) = k' \det_{\mathcal{B}}(v_1, \ldots, v_n)$ pour une certaine constante k'. En multipliant les deux membres de la relation $(*)$ par k', cette dernière devient alors pour la *même* constante k :

$$\det_{\mathcal{B}'}(f(v_1), \ldots, f(v_n)) = k \det_{\mathcal{B}'}(v_1, \ldots, v_n).$$

Ainsi, cette constante k ne dépend que de l'endomorphisme f. Par ailleurs, en remplaçant $v_1, \ldots, v_n$ dans $(*)$ par les vecteurs de la base $\mathcal{B}$, on trouve :

$$k = \det_{\mathcal{B}}(f(e_1), \ldots, f(e_n)), \tag{$**$}$$

de sorte que cette constante k est unique pour un endomorphisme f fixé. Cela justifie amplement la définition suivante :

Définition Soit E un $\mathbf{K}$-espace vectoriel de dimension n et f un endomorphisme de E. On appelle *déterminant de f* et l'on note $\det f$ l'unique scalaire tel que pour toute base $\mathcal{B}$ de E et tous vecteurs $v_1, \ldots, v_n \in E$:

$$\det_{\mathcal{B}}(f(v_1), \ldots, f(v_n)) = \det f\ \det_{\mathcal{B}}(v_1, \ldots, v_n).$$

Proposition 9 *Soit E un* **K***-espace vectoriel de dimension finie et f, g deux endomorphismes de E. Alors :*
$$\det(f \circ g) = \det f \, \det g.$$

Démonstration Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. D'après la relation $(**)$ ci-dessus, on a :

$$\det(f \circ g) = \det_{\mathcal{B}}\big((f \circ g)(e_1), \ldots, (f \circ g)(e_n)\big),$$

$$\det g = \det_{\mathcal{B}}\big(g(e_1), \ldots, g(e_n)\big).$$

Par définition de $\det f$, il en résulte :

$$\det(f \circ g) = \det_{\mathcal{B}}\big(f(g(e_1)), \ldots, f(g(e_n))\big) = \det f \, \det_{\mathcal{B}}\big(g(e_1), \ldots, g(e_n)\big) = \det f \, \det g.$$

□

Calcul du déterminant d'un endomorphisme

Soit E un **K**-espace vectoriel de dimension finie et f un endomorphisme de E. *Alors* $\det f$ *est le déterminant d'une matrice de f dans une base quelconque de E.* En effet, si A est la matrice de f dans une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E, alors les colonnes de A expriment les coordonnées dans $\mathcal{B}$ des vecteurs $f(e_1), \ldots, f(e_n)$. Il s'ensuit $\det_{\mathcal{B}}(f(e_1), \ldots, f(e_n)) = \det A$, d'où par la relation $(**)$ ci-dessus :
$$\det f = \det A.$$

Remarquons que deux matrices $n \times n$ quelconques A et B peuvent toujours être vues comme les matrices de deux endomorphismes f, g de $\mathbf{K}^n$ dans sa base canonique. La matrice produit AB est alors la matrice de l'endomorphisme $f \circ g$, de sorte que la proposition 9 peut être reformulée de la façon suivante :

Proposition 9 (bis) *Pour toutes matrices $n \times n$ A et B, on a :* $\det(AB) = \det A \, \det B$.

Remarque Deux matrices A, A' représentant un même endomorphisme f dans deux bases distinctes $\mathcal{B}$ et $\mathcal{B}'$ ont même déterminant, puisqu'il s'agit de l'unique déterminant $\det f$ de f. La proposition 10 permet de justifier directement ce fait ; on a alors $A = PA'P^{-1}$, où P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, et donc :

$$\det A = \det(PA'P^{-1}) = \det P \, \det(A'P^{-1}) = \det(A'P^{-1}) \det P = \det(A'P^{-1}P) = \det A'.$$

3.2 Critère d'inversibilité

Rappelons que pour un endomorphisme f d'un **K**-espace vectoriel E de dimension finie, les hypothèses suivantes sont équivalentes :

• f est injectif	• $\operatorname{Ker} f = \{0\}$	• $\dim \operatorname{Ker} f = 0$
• f est surjectif	• $\operatorname{Im} f = E$	• $\operatorname{rang} f \overset{\text{déf}}{=} \dim \operatorname{Im} f = \dim E$

Rappelons aussi que l'équivalence des hypothèses d'une même ligne est un fait général des applications linéaires (simple à établir) et que l'équivalence entre les deux lignes provient de la relation :

$$\dim \operatorname{Ker} f \; + \; \dim \operatorname{Im} f \; = \; \dim E.$$

Lorsque ces hypothèses équivalentes sont satisfaites, f est donc bijective — on dit aussi “inversible”, puisqu'alors la réciproque f^{-1} est bien définie et vérifie :

$$f \circ f^{-1} = f^{-1} \circ f = \mathrm{Id}_E.$$

Nous allons rallonger cette longue liste d'une hypothèse équivalente supplémentaire, en paraphrasant la proposition 4. Si $\mathcal{B} = (e_1, \ldots, e_n)$ est une base de E, alors dire que le rang de f est égal à la dimension de E revient à dire que les vecteurs $f(e_1), \ldots, f(e_n)$ sont linéairement indépendants. Selon la proposition 4, cela équivaut à l'hypothèse : $\det_{\mathcal{B}}(f(e_1), \ldots, f(e_n)) \neq 0$ ou encore, en vertu de la relation $(**)$ précédente :

Proposition 10 *Pour tout endomorphisme d'un* $\mathbf{K}$*-espace vectoriel de dimension finie :*

$$f \text{ est inversible} \quad \Longleftrightarrow \quad \det f \neq 0.$$

Rappelons que si A est la matrice d'un endomorphisme f dans une base donnée, alors sa matrice inverse A^{-1} — si elle existe — ne peut être que la matrice de l'application réciproque f^{-1}. Puisque toute matrice peut être vue comme celle d'un endomorphisme de $\mathbf{K}^n$ dans sa base canonique, la proposition 10 peut être reformulée, à l'instar de la proposition 9, de la façon suivante :

Proposition 10 (bis) *Pour toute matrice carrée A :*

$$A \text{ est inversible} \quad \Longleftrightarrow \quad \det A \neq 0.$$

Enfin, comme pour toute matrice $n \times n$ inversible A, A^{-1} est l'unique matrice B telle que : $AB = I_n$, le lemme 6 entraîne immédiatement :

Proposition 11 *Pour toute matrice inversible A, on a :*

$$A^{-1} = \frac{1}{\det A}\ {}^t\mathrm{Com}\, A.$$

Attention, cette jolie formule est, comme la méthode de Cramer, une calamité pour les calculs. Toute méthode d'inversion vue en 1$^{\text{ère}}$ année lui est préférable.

3.3 Valeurs propres et polynôme caractéristique d'un endomorphisme

En mécanique quantique, l'état d'un système est décrit par un vecteur unitaire v d'un $\mathbb{C}$-espace vectoriel E. Une grandeur physique observable (telle que la position en abscisse d'un photon, sa longueur d'onde...) y est formalisée par un endomorphisme f de E. Le système n'est observable que dans un état *propre*, c'est-à-dire un état v tel que $f(v)$ est parallèle à v. Le formalisme quantique ne détermine pas l'état du système, mais seulement les probabilités respectives — lors d'une mesure — des différents états propres. La grandeur physique effectivement mesurée est alors le coefficient de proportionnalité entre v et $f(v)$, autrement dit le scalaire λ tel que $f(v) = \lambda v$, appelé *valeur propre associée à l'état propre v*.

Nous allons nous intéresser ici au problème de la détermination des *valeurs propres d'un endomorphisme f* dans le cas le plus simple (insuffisant pour les besoins de la mécanique quantique) : celui où f est un endomorphisme d'un espace *E de dimension finie*.

Commençons par quelques définitions formelles.

Définition Soit f un endomorphisme d'un $\mathbf{K}$-espace vectoriel E.

- On dit qu'un scalaire $\lambda \in \mathbf{K}$ est *valeur propre* de f s'il existe un vecteur v non nul tel que :
$$f(v) = \lambda\, v.$$
- Ce vecteur v est alors appelé *vecteur propre* de f pour la valeur propre λ.
- L'ensemble des valeurs propres de f est appelé *spectre de f* et noté $\mathrm{Spec}(f)$.[1]

$\lambda \in \mathbf{K}$ est donc valeur propre d'un endomorphisme f si et seulement s'il existe un vecteur non nul v tel que : $(f - \lambda\,\mathrm{Id}_E)(v) \overset{\text{déf}}{=} f(v) - \lambda\, v = 0$, autrement dit ssi l'endormorphisme $f - \lambda\,\mathrm{Id}_E$ est non inversible, soit encore, d'après la proposition 10, si et seulement si $\det(f - \lambda\,\mathrm{Id}_E) = 0$. Les valeurs propres d'un endomorphisme f sont donc les solutions de l'équation :

$$\det(f - x\,\mathrm{Id}_E) = 0.$$

Ce fait nous invite à nous intéresser à la fonction : $x \mapsto \det(f - x\,\mathrm{Id}_E)$.

Proposition 12 *Soit E un $\mathbf{K}$-espace vectoriel de dimension finie n, f un endomorphisme de E et $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$ la matrice de f dans une base quelconque de E. L'application $P_f : \mathbf{K} \to \mathbf{K}$ définie par :*

$$P_f(x) = \det(f - x\,\mathrm{Id}_E)$$

est une fonction polynômiale de la forme :

$$P_f(x) = (-1)^n x^n + (-1)^{n-1}(a_{11} + \cdots + a_{nn})x^{n-1} + \cdots\cdots + \det A,$$

appelée polynôme caractéristique *de f. Les valeurs propres de f sont exactement les racines de P_f ; en particulier, f a* au plus *n valeurs propres.*

Démonstration Commençons par vérifier par récurrence sur $n \geqslant 1$ que pour toutes matrices $n\times n$ $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$ et $B = (b_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$, l'application $P : x \mapsto \det(A - xB)$ est une fonction polynômiale de degré au plus n :

- C'est évident lorsque $n = 1$, car alors : $\det(A - xB) = -b_{11}x + a_{11}$.
- Supposons cette propriété établie au rang $n-1$. En développant $\det(A-xB)$ selon la première colonne, on obtient : $\det(A - xB) = \sum_{i=1}^{n}(-1)^{i+1}(a_{i1} - xb_{i1})\det(A_{i1} - xB_{i1})$. Par hypothèse de récurrence, les applications $P_i : x \mapsto \det(A_{i1} - xB_{i1})$ sont des fonctions polynômiales de degrés $\leqslant n-1$. Ainsi, $\det(A - xB) = \sum_{i=1}^{n}(-1)^{i+1}(-b_{i1}x + a_{i1})P_i(x)$ est bien un polynôme de degré $\leqslant n$.

Si de plus tous les coefficients de la première ligne de B sont nuls, alors $P(x) = \det(A - xB)$ est un polynôme de degré $< n$. En effet :

- C'est évident lorsque $n = 1$, car alors $b_{11} = 0$, d'où : $\det(A - xB) = a_{11}$.
- Si $n \geqslant 2$, il suffit pour s'en convaincre de développer $\det(A - xB)$ selon la première ligne : $\det(A-xB) = \sum_{j=1}^{n}(-1)^{j+1}a_{1j}\det(A_{1j} - xB_{1j})$; or les applications $P_i : x \mapsto \det(A_{1j} - xB_{1j})$ sont polynômiales de degrés $\leqslant n-1$, en vertu de ce que l'on vient d'établir.

[1] Cette terminologie et la notation λ des valeurs propres sont une survivance d'un travail pionnier de Heisenberg, où les valeurs propres représentaient les longueurs d'ondes de photons émis par un atome d'hydrogène.

Établissons maintenant par récurrence sur $n \geqslant 1$ que pour toute matrice $n \times n$ $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ l'application $P : x \mapsto \det(A - xI_n)$ est une fonction polynômiale de la forme :

$$P(x) = (-1)^n x^n + (-1)^{n-1}(a_{11} + \cdots + a_{nn})x^{n-1} + \cdots\cdots + \det A.$$

- Lorsque $n = 1$, $P(x) = -x + a_{11}$ (dans ce cas, les monômes $(-1)^{n-1}(a_{11} + \cdots + a_{nn})x^{n-1}$ et $\det A$ de la forme générale représentent le même terme constant a_{11}).
- Prenons $n \geqslant 2$ et supposons la propriété établie au rang $n-1$. En développant $\det(A - xI_n)$ selon la première colonne, on obtient :

  $$\begin{aligned}\det(A - xI_n) &= (a_{11} - x)\det(A - xI_n)_{11} + \textstyle\sum_{i=2}^n (-1)^{i+1} a_{i1} \det(A_{i1} - xI_n)_{i1}\\ &= (a_{11} - x)\det(A_{11} - xI_{n-1}) + \textstyle\sum_{i=2}^n (-1)^{i+1} a_{i1} \det(A_{i1} - x(I_n)_{i1})\end{aligned}$$

  Pour tout $i \in \{2, \ldots, n\}$, la première ligne de la matrice $(n-1) \times (n-1)$ $(I_n)_{i1}$ ne contient que des coefficients nuls ; l'application $P_i : x \mapsto \det(A_{i1} - x(I_n)_{i1})$ est donc une fonction polynômiale de degré $< n-1$, d'après ce que l'on a établi plus haut. De plus, par hypothèse de récurrence, l'application $P_1 : x \mapsto \det(A_{11} - xI_{n-1})$ est un polynôme de la forme :

  $$P_1(x) = (-1)^{n-1}x^{n-1} + (-1)^n(a_{22} + \cdots + a_{nn})x^{n-2} + \cdots\cdots + \det B.$$

  En résumé, on a :

  $$\begin{aligned}P(x) &= (a_{11} - x)P_1(x) + \textstyle\sum_{i=2}^n (-1)^{i+1} a_{i1} P_i(x)\\ &= a_{11}P_1(x) - xP_1(x) + \textstyle\sum_{i=2}^n (-1)^{i+1} a_{i1} P_i(x)\\ &= a_{11}(-1)^{n-1}x^{n-1} - x((-1)^{n-1}x^{n-1} + (-1)^n(a_{22} + \cdots + a_{nn})x^{n-2}) + R(x)\\ &= a_{11}(-1)^{n-1}x^{n-1} + (-1)^n x^n + (-1)^{n-1}(a_{22} + \cdots + a_{nn})x^{n-1} + R(x)\\ &= (-1)^n x^n + (-1)^{n-1}(a_{11} + \cdots + a_{nn})x^{n-1} + R(x)\end{aligned}$$

  où $R(x)$ est un polynome de degré $\leqslant n-2$, autrement dit :

  $$P(x) = (-1)^n x^n + (-1)^{n-1}(a_{11} + \cdots + a_{nn})x^{n-1} + \cdots\cdots + c_1 x + c_0.$$

  En faisant $x = 0$, on obtient : $P(0) = c_0 = \det A$; $P(x)$ est bien de la forme annoncée.

Soit enfin f un endomorphisme f d'un **K**-espace vectoriel E de dimension n. Si une matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ représente f dans une base quelconque de E, alors l'endomorphisme $f - x\,\mathrm{Id}_E$ est représenté dans la même base par la matrice $A - xI_n$, d'où : $P_f(x) = \det(f - x\,\mathrm{Id}_E) = \det(A - xI_n)$ et la proposition s'ensuit.

$\square$

Remarque Le coefficient de x^{n-1} dans le polynôme P_f est $(-1)^{n-1}\sum_{i=1}^n a_{ii}$. Le scalaire :

$$\mathrm{tr}(A) = \sum_{i=1}^n a_{ii}$$

s'appelle la *trace* de la matrice A. Puisque le polynôme caractéristique ne dépend pas de la base choisie pour représenter f, il en est de même pour ses coefficients : cela signifie que les matrices représentant f dans les différentes bases de E ont toutes la même trace ; cette dernière est appelée *trace de l'endomorphisme* f et notée : $\mathrm{tr}\, f$.

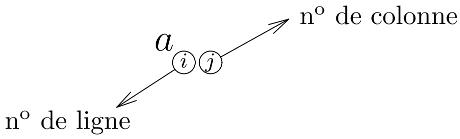
Annexe : Le déterminant d'une matrice en une formule

Dans la section 2.2 Existence et calcul (p. 14), nous avons exprimé pour $n = 1, 2, 3$ le déterminant d'une matrice $(a_{ij})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}}$ quelconque directement à partir de ses n^2 coefficients a_{ij}. Chaque fois, le déterminant s'est révélé être un polynôme des n^2 variables a_{ij}. L'objet de cette annexe est de préciser *pour tout entier* n la forme de ce polynôme. On aboutira ainsi à une définition synthétique des déterminants qui démystifiera quelque peu le sujet.

Commençons par présenter informellement la chose. Soit la matrice $n\times n$:

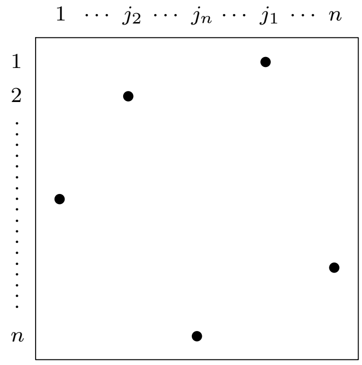
$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{1n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Rappel :

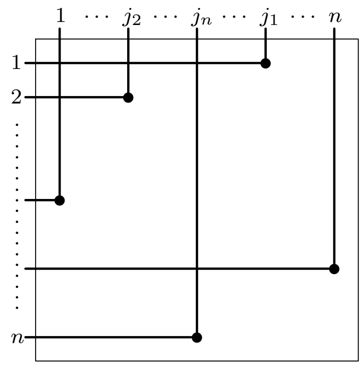


det A est alors une vaste somme $\sum_k \pm p_k$, où chaque p_k est le produit de n coefficients de la matrice A pris dans des lignes distinctes et des colonnes distinctes. À chaque tel choix de n coefficients de A correspond un terme $\pm p_k$ de cette somme. Comme il y a exactement $n!$ façons de choisir n coefficients dans des lignes et des colonnes distinctes, la somme $\sum_k \pm p_k$ comporte donc $n!$ termes.

La détermination du signe $\pm$ devant chaque p_k est un peu plus délicate. À partir des n coefficients choisis pour le produit p_k, mettons $p_k = a_{1j_1}a_{2j_2}\dots a_{nj_n}$, soit encore graphiquement :



on réalise le cablage suivant :

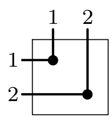
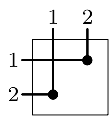


puis on compte le nombre total N d'intersections entre les fils du cablage. Le produit p_k est alors précédé du signe $+$ si N est pair, du signe $-$ si N est impair.

Afin d'éclairer cette définition informelle de déterminant, commençons par vérifier qu'elle redonne bien pour $n = 1, 2, 3$ les expressions trouvées p. 14.

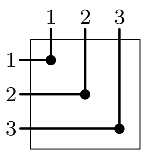
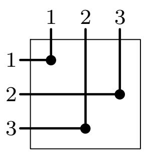
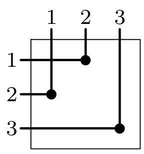
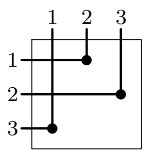
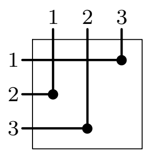
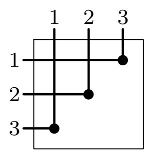
Lorsque $n = 1$, il n'y a évidemment qu'un seul choix possible de $n = 1$ coefficient dans la matrice A ; le nombre d'intersections de fils dans le cablage associé à cet unique choix a_{11} est $N = 0$ (puisque ce cablage ne comporte qu'un seul fil !), d'où : $\det A = +a_{11}$.

Lorsque $n = 2$, il n'y a que deux manières possibles de choisir $n = 2$ coefficients dans des lignes et des colonnes différentes :

choix de coefficients	a_{11}, a_{22}	a_{12}, a_{21}
cablage associé	 $N = 0$	 $N = 1$
terme	$+a_{11}a_{22}$	$-a_{12}a_{21}$

d'où : $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = +a_{11}a_{22} - a_{12}a_{21}$.

Lorsque $n = 3$, il y a exactement $3! = 6$ façons de choisir 3 coefficients dans des lignes et des colonnes différentes de A :

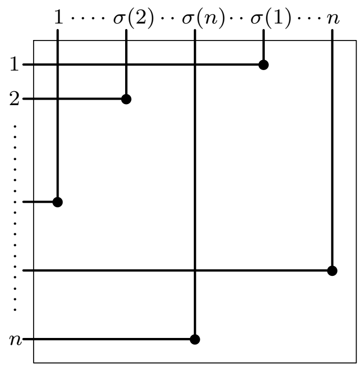
choix de coeff.	a_{11}, a_{22}, a_{33}	a_{11}, a_{23}, a_{32}	a_{12}, a_{21}, a_{33}	a_{12}, a_{23}, a_{31}	a_{13}, a_{21}, a_{32}	a_{13}, a_{22}, a_{31}
cablage associé	 $N = 0$	 $N = 1$	 $N = 1$	 $N = 2$	 $N = 2$	 $N = 3$
terme	$+a_{11}a_{22}a_{33}$	$-a_{11}a_{23}a_{32}$	$-a_{12}a_{21}a_{33}$	$+a_{12}a_{23}a_{31}$	$+a_{13}a_{21}a_{32}$	$-a_{13}a_{22}a_{31}$

d'où : $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = +a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$.

Nous allons maintenant résumer la définition de déterminant donnée ci-dessus en une formule. Le choix de coefficients correspondant à un terme quelconque $\pm p_k$ du déterminant se laisse décrire par l'application $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ qui à tout $i \in \{1, 2, \ldots, n\}$ associe le numéro de colonne du coefficient choisi dans la ligne n°i. Pour tout choix de n coefficients décrit de la sorte, il n'y a évidemment qu'un seul coefficient par ligne. Une telle application σ décrit donc un choix correct si et seulement si :

$$\begin{array}{ll} & \text{pour tout } j \in \{1, 2, \ldots, n\}, \text{ un seul coefficient a été choisi dans la colonne n}^\circ j \\ \Leftrightarrow & \text{pour tout } j \in \{1, 2, \ldots, n\}, \text{ il existe un seul } i \in \{1, 2, \ldots, n\} \text{ tel que } \sigma(i) = j \\ \Leftrightarrow & \text{l'application } \sigma \text{ est une bijection.} \end{array}$$

Ainsi les $n!$ choix corrects de n coefficients dans la matrice A sont décrits par les $n!$ bijections de $\{1, \ldots, n\}$ sur $\{1, \ldots, n\}$. Chacune de ces bijections $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ est naturellement représentée par un cablage comme précédemment :



Définition

(i). On appelle *permutation de* $\{1, \ldots, n\}$ une bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$.

(ii). On note $\mathfrak{S}_n$ l'ensemble des $n!$ permutations de $\{1, \ldots, n\}$.

(iii). On appelle *signature* de $\sigma \in \mathfrak{S}_n$ et l'on note $\mathrm{sgn}(\sigma)$, l'entier :

$$\mathrm{sgn}(\sigma) \overset{\text{déf}}{=} (-1)^N,$$

où N désigne le nombre d'intersections de fils dans le cablage représentant σ.

Il résulte de cette définition que dans un terme $\pm p_k$ du déterminant de A correspondant à un choix de coefficients décrit par $\sigma \in \mathfrak{S}_n$, le signe $\pm$ est $+$ si $\mathrm{sgn}(\sigma) = 1$, $-$ si $\mathrm{sgn}(\sigma) = -1$. Ainsi, on a : $\pm p_k = \mathrm{sgn}(\sigma)\, p_k$. Par ailleurs, $p_k = a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$, donc le déterminant de la matrice A s'exprime concisément comme suit.

$$\boxed{\det A \;=\; \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma)\, a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}}$$

Maintenant que nous avons obtenu une expression concise de la "définition" de déterminant présentée dans cette annexe, il est plus que temps de vérifier qu'elle est correcte, autrement dit qu'elle coïncide bien avec celle donnée plus tôt (cf section 2). Cette vérification utilise une propriété générale des applications multilinéaires alternées que l'on établit à l'aide de la proposition suivante.

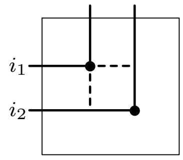
Proposition 13 *Soit σ une permutation de $\{1, \ldots, n\}$ et τ la permutation obtenue en échangeant deux valeurs quelconques $\sigma(i_1), \sigma(i_2)$ de σ $(i_1 \neq i_2)$. Autrement dit, τ est définie par :*

- $\tau(i_1) = \sigma(i_2)$,
- $\tau(i_2) = \sigma(i_1)$,
- $\tau(i) = \sigma(i)$ *si* $i \notin \{i_1, i_2\}$.

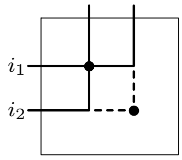
On a alors :

$$\mathrm{sgn}(\tau) = -\mathrm{sgn}(\sigma).$$

Démonstration Les cablages représentant les permutations σ et τ ne diffèrent que par leurs fils n°i_1 et n°i_2 :



$$N = L + \sum_{i \in I} N_i$$



$$N' = L + \sum_{i \in I} N'_i + 1$$

Le nombre L d'intersections de fils distincts des fils n°i_1 et n°i_2 est donc identique dans les deux cablages. La relation mentionnée sous chacun d'eux décrit le nombre total N (resp. N') d'intersections dans le cablage. I y désigne l'ensemble des n$^{\text{os}}$ de fils autres que i_1, i_2 et pour tout $i \in I$, N_i (resp. N'_i) dénote le nombre d'intersections du fil n°i avec les fils n°i_1 et n°i_2.

Clairement, si le fil n°i débute et termine entre les fils n°i_1 et n°i_2, alors $N_i = 0$, $N'_i = 2$. Dans tous les autres cas, on vérifie facilement la relation $N_i = N'_i$. Ainsi, pour tout $i \in I$, $N'_i - N_i$ est pair, donc $N' - N = \sum_{i \in I}(N'_i - N_i) + 1$ est impair. Il s'ensuit : $(-1)^{N'}/(-1)^N = (-1)^{N'-N} = -1$; autrement dit, les signatures de σ et de τ sont opposées.

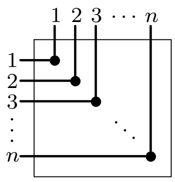
□

Proposition 14 *Soit $\varphi : E^n \to \mathbf{K}$ une application n-linéaire alternée sur un $\mathbf{K}$-espace vectoriel E quelconque ($\mathbf{K} = \mathbb{R}$ ou $\mathbb{C}$) et σ une permutation de $\{1, 2, \ldots, n\}$. Alors pour tous vecteurs $v_1, v_2, \ldots, v_n$ de E, on a :*

$$\varphi(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)}) = \operatorname{sgn}(\sigma)\, \varphi(v_1, v_2, \ldots, v_n).$$

Démonstration Montrons cette relation par récurrence sur le nombre p d'entiers $i \in \{1, 2, \ldots, n\}$ tels que $\sigma(i) \neq i$.

Lorsque $p = 0$, on a $\sigma(i) = i$ pour tout $i \in \{1, 2, \ldots, n\}$; le cablage représentant σ n'a alors aucune intersection de fil :



$$N = 0$$

d'où : $\operatorname{sgn}(\sigma) = (-1)^0 = 1$ et la relation s'ensuit immédiatement.

Supposons maintenant que le nombre d'entiers $i \in \{1, 2, \ldots, n\}$ tels que $\sigma(i) \neq i$ vaut $p \geqslant 1$ et que la relation est vérifiée pour tout nombre inférieur. Comme $p \neq 0$, il existe $i_1 \in \{1, 2, \ldots, n\}$ tel que $\sigma(i_1) \neq i_1$. L'antécédent i_2 de i_1 par σ est donc distinct de i_1 : $\sigma(i_2) = i_1 \neq i_2$. Soit τ la permutation obtenue en échangeant les valeurs $\sigma(i_1), \sigma(i_2)$ de σ comme dans la proposition précédente. On a par conséquent : $\tau(i_1) = \sigma(i_2) = i_1$ et le nombre d'entiers $i \in \{1, 2, \ldots, n\}$ tels que $\tau(i) \neq i$ vaut $p - 1$ si $\tau(i_2) \neq i_2$, $p - 2$ sinon. L'hypothèse de récurrence entraîne donc : $\varphi(v_{\tau(1)}, v_{\tau(2)}, \ldots, v_{\tau(n)}) = \operatorname{sgn}(\tau)\, \varphi(v_1, v_2, \ldots, v_n)$. Comme selon la proposition 13, on a $\operatorname{sgn}(\tau) = -\operatorname{sgn}(\sigma)$, le lemme 1 (ii) p. 7 permet d'en déduire :

$$\begin{aligned}\varphi(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) &= -\varphi(v_{\sigma(1)}, \ldots, v_{\sigma(i_1-1)}, v_{\sigma(i_2)}, v_{\sigma(i_1+1)}, \ldots, v_{\sigma(i_2-1)}, v_{\sigma(i_1)}, v_{\sigma(i_2+1)}, \ldots, v_{\sigma(n)}) \\ &= -\varphi(v_{\tau(1)}, v_{\tau(2)}, \ldots, v_{\tau(n)}) = -\operatorname{sgn}(\tau)\, \varphi(v_1, \ldots, v_n) = \operatorname{sgn}(\sigma)\, \varphi(v_1, \ldots, v_n).\end{aligned}$$

□

Cette dernière proposition nous permet de calculer un déterminant quelconque comme suit.

Théorème 15 *Pour toute matrice à coefficients dans* $\mathbf{K}$ $(\mathbf{K} = \mathbb{R}$ *ou* $\mathbb{C})$ *:*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{1,n} \\ \vdots & & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix},$$

on a :

$$\det A \;=\; \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma)\, a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)}.$$

Démonstration $\det A$ désigne évidemment ici le déterminant de A tel que défini dans la section 2. Par définition, $\det A = \det_{\mathcal{B}}(u_1, \ldots, u_n)$, où $\mathcal{B} = (e_1, \ldots, e_n)$ est une base quelconque de n'importe quel $\mathbf{K}$-espace vectoriel E de dimension n et où pour tout $j \in \{1, \ldots, n\}$, u_j est le vecteur de coordonnées $\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$ dans la base $\mathcal{B} = (e_1, \ldots, e_n)$.

Comme en vertu de la proposition 7 (p. 17), $\det A = \det {}^tA$, on a aussi : $\det A = \det_{\mathcal{B}}(v_1, \ldots, v_n)$, où pour tout $i \in \{1, \ldots, n\}$, v_i est le vecteur de coordonnées $\begin{pmatrix} a_{i,1} & \cdots & a_{i,n} \end{pmatrix}$ dans la base $\mathcal{B}$, autrement dit $v_i = \sum_{j=1}^{n} a_{i,j} e_j$.

En utilisant successivement la linéarité de l'application $\det_{\mathcal{B}}$ par rapport à chacun de ses arguments, on obtient :

$$\begin{aligned} \det A &= \det_{\mathcal{B}}\Big(\sum_{j_1=1}^{n} a_{1,j_1} e_{j_1}, \sum_{j_2=1}^{n} a_{2,j_2} e_{j_2}, \ldots, \sum_{j_n=1}^{n} a_{n,j_n} e_{j_n} \Big) \\ &= \sum_{j_1=1}^{n} a_{1,j_1} \det_{\mathcal{B}}\Big(e_{j_1}, \sum_{j_2=1}^{n} a_{2,j_2} e_{j_2}, \ldots, \sum_{j_n=1}^{n} a_{n,j_n} e_{j_n} \Big) \\ &= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} a_{1,j_1} a_{2,j_2} \det_{\mathcal{B}}\Big(e_{j_1}, e_{j_2}, \sum_{j_3=1}^{n} a_{3,j_3} e_{j_3}, \ldots, \sum_{j_n=1}^{n} a_{n,j_n} e_{j_n} \Big) \\ &\;\vdots \\ &= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1,j_1} a_{2,j_2} \ldots a_{n,j_n} \det_{\mathcal{B}}\big(e_{j_1}, e_{j_2}, \ldots, e_{j_n} \big). \end{aligned}$$

Cette somme de n sommes de … de n sommes n'est jamais qu'une somme de $\overbrace{n \times n \times \ldots \times n}^{n \text{ facteurs}} = n^n$ termes dont chacun correspond à l'un des n^n choix d'entiers $j_1, \ldots, j_n$ entre 1 et n. Chacun de ces choix d'entiers $j_1, \ldots, j_n$ peut être représenté par l'application $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ qui à tout i associe j_i, de sorte que notre somme de sommes de … de sommes s'écrit aussi :

$$\det A \;=\; \sum_{f \in \mathcal{F}} a_{1,f(1)} a_{2,f(2)} \ldots a_{n,f(n)} \det_{\mathcal{B}}\big(e_{f(1)}, e_{f(2)}, \ldots, e_{f(n)} \big),$$

où $\mathcal{F}$ désigne l'ensemble des n^n applications de $\{1, 2, \ldots, n\}$ dans $\{1, 2, \ldots, n\}$. Cependant, lorsque f n'est pas une bijection, il existe nécessairement deux entiers $i_1 \neq i_2$ tels que $f(i_1) = f(i_2)$ et donc $e_{f(i_1)} = e_{f(i_2)}$; comme l'application $\det_{\mathcal{B}}$ est alternée, il en résulte que le terme correspondant est nul. On peut donc restreindre notre somme aux autres termes :

$$\det A \;=\; \sum_{\sigma \in \mathfrak{S}_n} a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)} \det_{\mathcal{B}}\big(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)} \big).$$

Enfin, puisque par définition $\det_{\mathcal{B}} : E^n \to \mathbf{K}$ est l'unique application n-linéaire alternée telle que $\det_{\mathcal{B}}(e_1, \ldots, e_n) = 1$ (cf définition au bas de la page 8), on obtient à l'aide de la proposition 14 :

$$\begin{aligned}
\det A &= \sum_{\sigma \in \mathfrak{S}_n} a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)} \det_{\mathcal{B}}\big(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}\big) \\
&= \sum_{\sigma \in \mathfrak{S}_n} a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)} \operatorname{sgn}(\sigma) \det_{\mathcal{B}}\big(e_1, e_2, \ldots, e_n\big) \\
&= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma)\, a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)}.
\end{aligned}$$

□

L'expression de $\det A$ que l'on vient d'établir en toute généralité possède la vertu d'éclairer, sinon rendre évident, un certain nombre de faits vus dans ce chapitre. Par exemple, puisque ${}^tA$ est la matrice “symétrique par rapport à la diagonale de A” :

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{1n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \qquad {}^tA = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \vdots & & \ddots & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix}$$

et puisque les $n!$ choix de n coefficients dans des lignes et des colonnes différentes se transforment globalement en eux-mêmes par cette “symétrie”, qui par ailleurs ne change pas le nombre d'intersections de fils dans les cablages associés, il devient tout à fait clair que l'on a : $\det A = \det {}^tA$ (cf proposition 7, p. 17).

Un autre fait facilement expliqué par cette formule est la forme générale du polynôme caractéristique d'un endomorphisme f (cf proposition 12, p. 23). Soit $A = (a_{ij})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}}$ la matrice de f dans une base quelconque ; le polynôme caractéristique de f est donc $\det(A - xI_n)$. Le terme de ce déterminant correspondant au choix des n coefficients diagonaux est $+(a_{11}-x)(a_{22}-x)\ldots(a_{nn}-x)$, car le cablage associé n'a aucune intersection de fil (voir le graphique de la démonstration de la proposition 14). Tout autre terme de $\det(A - xI_n)$ possède au moins un coefficient non diagonal a_{ij} $(i \neq j)$ et ne peut donc contenir les coefficients diagonaux $a_{ii} - x$, $a_{jj} - x$ situés sur la même ligne (resp. la même colonne) ; ainsi, tout autre terme contient au plus $n-2$ coefficients diagonaux et constitue donc un polynôme de degré au plus $n - 2$. Il s'ensuit que le polynôme caractéristique de f est de la forme :

$$P_f(x) = (a_{11} - x)(a_{22} - x)\ldots(a_{nn} - x) + Q(x),$$

où Q est de degré au plus $n - 2$. Seul le premier terme de cette expression contribue donc aux termes en x^n et en x^{n-1} de $P_f(x)$; en le développant, on obtient donc :

$$P_f(x) = (-1)^n x^n + (-1)^{n-1}(a_{11} + \cdots + a_{nn})x^{n-1} + \cdots\cdots + c_1 x + c_0.$$

Enfin, en faisant $x = 0$, on obtient : $P_f(0) = c_0 = \det A$, d'où la forme :

$$P_f(x) = (-1)^n x^n + (-1)^{n-1}(a_{11} + \cdots + a_{nn})x^{n-1} + \cdots\cdots + \det A.$$