Formes réduites échelonnées

Antoine Chambert-Loir

1. Opérations élémentaires sur les matrices

Soit K un anneau, non nécessairement commutatif. On considère K^n comme un K-module à droite, de sorte que les applications linéaires de K^p dans K^n s'identifient aux matrices de $\mathrm{Mat}_{n,p}(K)$ à n lignes et p colonnes, agissant par les formules habituelles. La composition des applications linéaires correspond à la multiplication des matrices.

1.1. *Matrices élémentaires.* — Le groupe $\mathrm{GL}_n(K)$ des matrices inversibles $n \times n$ à coefficients dans K contient un certain nombre d'éléments importants.

On note $(e_{i,j})$ la base canonique de $\mathrm{Mat}_n(K)$; $e_{i,j}$ est la matrice dont tous les coefficients sont nuls sauf celui de la ligne i et de la colonne j qui vaut 1. Pour $i, j \in \{1, \dots, n\}$, $i \neq j$, et $a \in K$, on pose $E_{ij}(a) = I_n + ae_{i,j}$, où I_n est la matrice identité. On a la relation

$$E_{ij}(a)E_{ij}(b) = E_{ij}(a+b)$$

qui, jointe à l'identité évidente $E_{ij}(0) = I_n$, entraîne que les matrices $E_{ij}(a)$ sont inversibles.

Si $\sigma \in \mathfrak{S}_n$, on note P_σ la matrice attachée à σ ; c'est celle de l'application canonique qui applique le j-ième vecteur de base sur le $\sigma(j)$-ième. Autrement dit, si $P_\sigma = (p_{i,j})$, on a $p_{i,j} = 1$ si $i = \sigma(j)$ et $p_{i,j} = 0$ sinon. On a $P_{\sigma\tau} = P_\sigma P_\tau$ et $P_{\mathrm{id}} = I_n$. L'application $\sigma \mapsto P_\sigma$ est un isomorphisme du groupe $\mathfrak{S}_n$ sur un sous-groupe de $\mathrm{GL}_n(K)$ que l'on note W.

Pour $1 \leqslant j \leqslant n$ et $a \in K$, on note enfin $D_j(a)$ la matrice diagonale $I_n + (a-1)e_{j,j}$ dont les coefficients diagonaux sont tous égaux à 1 sauf celui de la ligne j et de la colonne j qui vaut a. On a $D_j(a)D_j(b) = D_j(ab)$ et $D_j(1) = I_n$; si $a \in K^*$, alors $D_j(a)$ appartient à $\mathrm{GL}_n(K)$.

On note $\mathrm{GE}_n(K)$ le sous-groupe de $\mathrm{GL}_n(K)$ engendré par les matrices élémentaires : les matrices $E_{i,j}(a)$, pour $a \in K$, les matrices de permutation P_σ et les matrices $D_j(a)$, pour $a \in K^*$.

1.2. *Opérations élémentaires.* — Soit M une matrice à n lignes et p colonnes à coefficients dans K.

La multiplication à droite de M par les matrices élémentaires (de $\mathrm{Mat}_p(K)$) correspond aux manipulations classiques sur les colonnes de M. La matrice $ME_{i,j}(a)$ est obtenue en ajoutant la i-ième colonne de M fois a à sa j-ième colonne (opération qu'on symbolise par $C_j \leftarrow C_j + C_i a$). La matrice MP_σ est obtenue en permutant les colonnes de M : la j-ième colonne de MP_σ est la $\sigma(j)$-ième de M. La matrice $MD_j(a)$ est obtenue en multipliant la j-ième colonne de M par a (soit encore $C_j \leftarrow C_j a$).

La multiplication à gauche de M par les matrices élémentaires (de $\mathrm{Mat}_n(K)$) correspond, quant à elle, aux opérations classiques sur les lignes de M. La matrice $E_{i,j}(a)M$ est obtenue en ajoutant a fois la j-ième ligne de M à sa i-ième ligne (on note $L_i \leftarrow L_i + aL_j$) ; la i-ième ligne de M est la ligne d'indice $\sigma(i)$ de la matrice $P_\sigma M$; les lignes de $D_i(a)M$ sont celles de M sauf la ligne d'indice i qui est multipliée par i (c'est-à-dire $L_i \leftarrow aL_i$).

1.3. *Équivalences par lignes, par colonnes.* — On dit que deux matrices M et M' dans $\mathrm{Mat}_{n,p}(K)$ sont équivalentes par lignes s'il existe une matrice $A \in \mathrm{GE}_n(K)$ telle que $M' = AM$; cela revient à dire qu'il existe une suite d'opérations élémentaires sur les lignes qui fait passer de M à M'. C'est une relation d'équivalence ; ses classes d'équivalence ne sont autres que les orbites de l'action de $\mathrm{GE}_n(K)$ par multiplication à gauche sur $\mathrm{Mat}_{n,p}(K)$. Cela signifie encore que les systèmes linéaires dans K^p définis par M et M' ont mêmes solutions.

On dit de même que deux telles matrices sont équivalentes par colonnes s'il existe une matrice $A \in \mathrm{GE}_p(K)$ telle que $M' = MA$, autrement dit s'il existe une suite d'opérations élémentaires sur les colonnes qui fait passer de M à M'. C'est encore une relation d'équivalence dans $\mathrm{Mat}_{n,p}(K)$ dont les classes d'équivalence sont les orbites de l'action de $\mathrm{GE}_p(K)$ par multiplication à droite sur $\mathrm{Mat}_{n,p}(K)$. Cela signifie qu'il existe un changement linéaire $X' = AX$ d'inconnues, donné par un produit A de matrices élémentaires, transformant le système $MX = 0$ en le système $M'X' = 0$.

2. Formes réduites échelonnées

2.1. *Définition.* — Soit K un anneau.

On dit qu'une matrice $M = (m_{i,j})$ à coefficients dans K, de taille $n \times p$, est sous forme réduite échelonnée par lignes s'il existe un entier $r \in \{0, \dots, n\}$ et des entiers $j_1, \dots, j_r$ de sorte que les conditions suivantes soient satisfaites :

(1) on a $1 \leqslant j_1 < \dots < j_r \leqslant p$;

(2) si $1 \leqslant i \leqslant r$ et $1 \leqslant j < j_i$, on a $m_{i,j} = 0$;

(3) pour tout $1 \leqslant i \leqslant r$, on a $m_{i,j_i} = 1$ et $m_{k,j_i} = 0$ si $k \in \{1, \dots, n\}$ est différent de i ;

(4) si $i > r$, alors $m_{i,j} = 0$ pour tout $j \in \{1, \dots, p\}$.

Les coefficients $m_{i,j_i} = 1$ sont les pivots, les entiers $j_1, \dots, j_r$ sont appelés indices de pivot. Les conditions ci-dessus expriment donc que le premier coefficient non nul des r premières lignes est un pivot ; que les indices de pivot vont en croissant ; qu'hormis ce pivot, les coefficients d'une colonne de pivot sont nuls ; que les lignes d'indice $> r$ sont nulles.

Comme l'application $i \mapsto j_i$ est strictement croissante, on a $r \leqslant p$.

Une observation importante est que si M est sous forme réduite échelonnée par lignes, la matrice déduite de M par extraction des k premières colonnes est encore sous forme réduite échelonnée par lignes.

On dit de manière analogue qu'une matrice $A = (m_{i,j})$ à coefficients dans K, de taille $n \times p$, est sous forme réduite échelonnée par colonnes si sa transposée est sous forme réduite échelonnée par lignes. Cela signifie qu'il existe un entier $s \in \{0, \dots, p\}$ et des entiers $i_1, \dots, i_s$ de sorte que les conditions suivantes soient satisfaites :

(1) on a $1 \leqslant i_1 < \cdots < i_s \leqslant n$;
(2) si $1 \leqslant j \leqslant s$ et $1 \leqslant i < i_j$, on a $m_{i,j} = 0$;
(3) pour tout j tel que $1 \leqslant j \leqslant s$, on a $m_{i_j,j} = 1$ et $m_{i_j,k} = 0$ si $k \in \{1, \dots, p\}$ est différent de j ;
(4) si $j > s$, alors $m_{i,j} = 0$ pour tout $i \in \{1, \dots, n\}$.

2.2. *Toute matrice est équivalente suivant les lignes à une matrice réduite échelonnée par lignes.* — Soit K un corps, voire un anneau à division. Nous allons démontrer par la méthode d'élimination dite de Gauss–Jordan l'assertion énoncée en titre de ce paragraphe.

Soit $M \in \mathrm{Mat}_{n,p}(K)$. Pour $k \in \{0, \dots, p\}$, notons M_k la matrice de $\mathrm{Mat}_{n,k}(K)$ déduite de M par extraction des k premières colonnes. Démontrons par récurrence sur k qu'il existe une matrice $A \in \mathrm{GE}_n(K)$ telle que AM_k soit sous forme réduite échelonnée par lignes.

Il n'y a rien à démontrer pour $k = 0$. Supposons que l'assertion soit vraie pour $k-1$ et soit $A \in \mathrm{GE}_n(K)$ tel que AM_{k-1} soit sous forme réduite échelonnée par lignes, d'indices de pivots $j_1, \dots, j_r$. Considérons alors la matrice AM_k ; ses $k-1$ premières colonnes sont celles de la matrice AM_{k-1}. Considérons les coefficients $m_{i,k}$ pour $i > r$ de la matrice AM_k ; s'ils sont tous nuls, la matrice AM_k est sous forme réduite échelonnée par lignes. Sinon, quitte à échanger la ligne d'indice $r+1$ avec une ligne d'indice $i > r+1$, on peut supposer que $m_{r+1,k} \neq 0$. Divisons la ligne d'indice $r+1$ par $m_{r+1,k}$; le coefficient $m_{r+1,k}$ est alors égal à 1. Il reste à remplacer, pour tout $i \in \{1, \dots, n\}$ distinct de $r+1$, la ligne L_i par la ligne $L_i - m_{i,k}L_{r+1}$ pour obtenir une matrice sous forme réduite échelonnée par lignes dont les indices de pivot sont les entiers $j_1, \dots, j_r, k$.

L'assertion est alors vérifiée pour $k = p$, ce qui est précisément le résultat annoncé.

On notera la différence avec l'algorithme classique d'élimination de Gauss qui se contente d'annuler les coefficients sous le pivot, ce qui fournit seulement une matrice dont les lignes sont échelonnées.

2.3. *Application aux systèmes linéaires.* — Soit $M \in \mathrm{Mat}_{n,p}(K)$ et soit M' une matrice réduite échelonnée par lignes qui est équivalente à M. Les systèmes $MX = 0$ et $M'X = 0$ sont donc équivalents. L'intérêt de la mise sous forme réduite échelonnée est que le système $M'X = 0$ est entièrement résolu : les variables correspondant aux indices de pivot sont exprimées en fonction des autres.

Cela s'applique aussi aux systèmes avec un second membre $MX = Y$ qu'on interprète comme le système linéaire dans K^{p+1} $MX + yY = 0$ d'inconnue (X, y), auquel on adjoint la condition $y = -1$. Soit $[M'Y']$ une matrice sous forme réduite échelonnée par lignes qui est équivalente par lignes à la matrice $[MY]$. Le système $MX = Y$ est ainsi équivalent au système $M'X = Y'$. Il y a deux possibilités : si la dernière colonne de $[M'Y']$ n'est pas celle d'un pivot, le système $M'X = Y'$ est résolu, les variables de pivot étant exprimées en fonction des coefficients de Y' et des autres variables. Dans le cas contraire où la dernière colonne de $[M'Y']$ est celle d'un pivot, le système $M'X = Y'$ n'a pas de solution.

Soit $A \in \mathrm{GE}_n(K)$ une matrice telle que $M' = AM$ soit sous forme réduite échelonnée par lignes, d'indices de pivot $j_1, \dots, j_r$. Pour $Y \in K^p$, le système $MX = Y$ équivaut au système $M'X = AY$. Considérons un vecteur colonne Y dont les entrées sont des variables $y_1, \dots, y_n$; les $n - r$ dernières lignes de AY s'intérprètent comme les conditions linéaires qui doivent être satisfaites par Y pour que $MX = Y$ possède une solution ; si elles le sont, les r premières lignes de l'égalité $M'X = AY$ déterminent les variables de pivot en termes des autres et des coefficients de Y.

Cela se généralise encore aux conjonctions de systèmes $MX = Y_1, \dots, MX = Y_q$. Il suffit de calculer une matrice $[M'Y'_1 \dots Y'_q]$ sous forme réduite échelonnée par lignes qui est équivalente à la matrice $[MY_1 \dots Y_q]$. Si tous ses indices de pivot sont $\leqslant p$, le système est alors résolu ; sinon, il est inconsistant.

2.4. *Unicité de la forme réduite échelonnée par lignes.* — Soit M une matrice de $\mathrm{Mat}_{n,p}(K)$; nous allons prouver qu'il existe une seule matrice réduite échelonnée par lignes qui est équivalente à M. Il suffit de démontrer que si M et M' sont équivalentes suivant les lignes, et sont toutes deux sous forme réduite échelonnée par lignes, alors $M = M'$.

Comme pour l'assertion d'existence, pour $k \in \{0, \dots, p\}$, nous notons M_k et M'_k les matrices de taille $n \times k$ obtenue par extraction des k premières colonnes de M et M' respectivement. Démontrons par récurrence sur k que $M_k = M'_k$.

L'assertion est vraie pour $k = 0$; supposons-la vraie pour k, c'est-à-dire $M_k = M'_k$, et démontrons-la pour $k + 1$. Notons $j_1, \dots, j_r$ les indices de pivot de M_k. Notons aussi Y et Y' les dernières colonnes de M_{k+1} et M'_{k+1}, de sorte que $M_{k+1} = [M_kY]$ et $M'_{k+1} = [M'_kY']$; on doit prouver que $Y = Y'$. Puisque les matrices M_{k+1} et M'_{k+1} sont équivalentes suivant les lignes, les systèmes linéaires (avec second membre) $M_kX = Y$ et $M'_kX = Y'$ sont équivalents. Comme ces matrices sont sous forme réduite échelonnée par lignes, l'analyse de ces systèmes est très simple.

Le système $M_kX = Y$ est inconsistant si et seulement si $k + 1$ est un indice de pivot de M ; cela entraîne que $k + 1$ est un indice de pivot de M si et seulement si c'en est un de M'. Dans ce cas, $Y = Y'$ est le vecteur colonne ayant tous ses coefficients nuls sauf celui de la ligne $r + 1$ qui vaut 1.

Sinon, si $k+1$ n'est pas un indice de pivot de M (ou de M'), les deux systèmes $M_kX = Y$ et $M'_kX = Y'$ ont des solutions ; ils ont les mêmes solutions car ils sont équivalents. Comme $M_k = M'_k$, il vient nécessairement $Y = Y'$.

Cette assertion d'unicité permet de définir les indices de pivot d'une matrice $M \in \mathrm{Mat}_{n,p}(K)$ comme ceux de l'unique matrice réduite échelonnée par lignes qui est équivalente à M. De même, on appelle rang des lignes de M le nombre de ces indices de pivot.

Si $q \leqslant p$, on appellera plus généralement q-forme réduite échelonnée de M une matrice AM, où $A \in \mathrm{GE}_n(K)$ est telle que AM_q soit la forme réduite échelonnée de la matrice M_q formée des q premières colonnes de M.

2.5. *Formes réduites échelonnées par colonnes.* — De manière analogue, les orbites de l'action à droite de $\mathrm{GL}_p(K)$ par multiplication à droite sur $\mathrm{Mat}_{n,p}(K)$ contiennent chacune une matrice réduite échelonnée par colonnes et une seule.

3. Application à l'algèbre linéaire

3.1. *Retour aux systèmes linéaires.* — Soit r le rang des lignes et $j_1, \dots, j_r$ les indices de pivot d'une matrice $M \in \mathrm{Mat}_{n,p}(K)$.

Le système linéaire $MX = 0$ possède $X = 0$ comme unique solution si et seulement si toutes les variables sont des variables de pivot, c'est-à-dire si et seulement si $r = p$ et $j_i = i$ pour $1 \leqslant i \leqslant p$. Cela entraîne en particulier que $n \geqslant p$. On retrouve ainsi qu'*un système linéaire homogène ayant strictement plus d'inconnues que d'équations possède une solution non nulle* ; de manière équivalente, si $p > n$, une application linéaire de K^p dans K^n n'est pas injective.

Soit par ailleurs $Y \in K^p$. Soit $[M'Y']$ la matrice sous forme réduite échelonnée par lignes équivalente à la matrice augmentée $[MY]$. La matrice M' est sous forme réduite échelonnée par lignes, et est équivalente à M. Les indices de pivot de la matrice augmentée $[MY]$ sont donc soit $j_1, \dots, j_r$, soit $j_1, \dots, j_r, p+1$. Dans le premier cas, le système $MX = Y$ est équivalent au système résolu $M'X' = Y'$ donc possède une solution ; dans le second cas, le système $MX = Y$ n'a pas de solution.

On en déduit qu'*une application linéaire de K^p dans K^n n'est pas surjective si $p < n$.* En effet, il est alors possible de trouver des vecteurs Y tels que les indices de pivot de la matrice $[MY]$ soient $j_1, \dots, j_r, p+1$; il suffit par exemple de choisir $A \in \mathrm{GE}_n(K)$ tel que $M' = AM$ et de poser $Y = A^{-1}Y'$, où $Y' = (0, \dots, 0, 1)$. De manière générale, on pourra trouver de tels vecteurs Y dès que $r < n$.

Plaçons nous maintenant dans le cas où $n = p$. On voit que les conditions suivantes sont toutes équivalentes :

(1) l'endomorphisme de K^n défini par M est injectif ;
(2) $r = n$;
(3) M est équivalente suivant les lignes à la matrice identité ;
(4) l'endomorphisme de K^n défini par M est surjectif ;
(5) l'endomorphisme de K^n défini par M est bijectif ;
(6) $M \in \mathrm{GL}_n(K)$.

Une conséquence de ce résultat est l'égalité $\mathrm{GE}_n(K) = \mathrm{GL}_n(K)$: *le groupe linéaire* $\mathrm{GL}_n(K)$ *est engendré par les matrices élémentaires.*

Une autre conséquence est que les orbites de l'action de $\mathrm{GL}_n(K)$ par multiplication à gauche sur $\mathrm{Mat}_{n,p}(K)$ contiennent chacune une matrice réduite échelonnée par lignes et une seule.

3.2. *Résolution algorithmique d'un problème standard d'algèbre linéaire.* — Soit $Y_1, \dots, Y_p$ des vecteurs de K^n ; soit V le sous-espace engendré par $Y_1, \dots, Y_p$. Un problème standard d'algèbre linéaire est de déterminer une base de V, ou un système d'équations linéaires définissant V.

Soit M la matrice $[Y_1 \dots Y_p]$; on interprète le système $MX = 0$ comme fournissant les relations de dépendance linéaire entre les vecteurs Y_i. Si l'on impose les conditions supplémentaires $x_{j+1} = \dots = x_p = 0$, cela revient à rechercher les relations de dépendance linéaire entre $Y_1, \dots, Y_j$. Notons $j_1, \dots, j_r$ les indices de pivot de la matrice M. On obtient alors que $Y_{j_1}, \dots, Y_{j_r}$ est une base de V.

Cela fournit un moyen algorithmique de déterminer une base du sous-espace V_p de K^n engendré par les vecteurs Y_i : calculer la forme réduite échelonnée par lignes de la matrice $[Y_1 \dots Y_p]$: une base de V_p est formée des vecteurs $Y_{j_1}, \dots, Y_{j_r}$ correspondant aux indices de pivot.

Calculons plus généralement une p-forme réduite échelonnée par lignes $[M'Y']$ de la matrice $[Y_1 \dots Y_p Y]$, où Y est un vecteur colonne de coefficients indéterminés $(y_1, \dots, y_n)$. Les $n - r$ dernières lignes de la colonne Y' fournissent un système d'équations linéaires définissant V_p.

Plus généralement, pour $j \in \{0, \dots, p\}$, notons V_j le sous-espace vectoriel de K^n engendré par $Y_1, \dots, Y_j$. On constate ainsi que pour tout entier $k \in \{0, \dots, r\}$, $Y_{j_1}, \dots, Y_{j_k}$ est une base de V_j si $j_k \leqslant j < j_{k+1}$ (avec les conventions $j_0 = 0$ et $j_{r+1} = n + 1$).

3.3. *Dimension d'un espace vectoriel.* — Soit V un sous K-espace vectoriel de K^n. Soit $(Y_1, \dots)$ une suite de vecteurs de V choisie de sorte que chacun soit linéairement indépendant des précédents. Pour tout entier m, le rang des lignes de la matrice $[Y_1, \dots, Y_m]$ est alors égal à m, ce qui impose $m \leqslant n$. Il existe donc un entier p et des vecteurs $Y_1, \dots, Y_p$ de V tels que tout vecteur de V soit combinaison linéaire des Y_i.

On a vu au paragraphe précédent qu'alors que si V est engendré par des vecteurs $Y_1, \dots, Y_p$, alors V possède une base, extraite de la famille $Y_1, \dots, Y_p$, et de cardinal égal au rang r des lignes de la matrice $[Y_1 \dots Y_p]$. En particulier, $r \leqslant \min(n, p)$. Par le choix d'une base de cardinal b, V s'identifie à K^b ; comme une application linéaire de K^b dans K^c n'est pas injective si $c < b$, et n'est pas surjective si $b < c$, toutes les bases de V sont de cardinal r. On peut donc définir la dimension de V comme le cardinal d'une quelconque de ses bases.

Les résultats ci-dessus entraînent que toutes les bases de V ont même cardinal, nombre entier qu'on appelle la dimension de V.

3.4. *Rang des lignes, vs. rang des colonnes.* — La mise sous forme réduite échelonnée par colonnes a aussi son intérêt. Soit en effet $M = [Y_1 \dots Y_p]$ une matrice de $\mathrm{Mat}_{n,p}(K)$, de forme réduite échelonnée par colonnes $M' = [Y'_1 \dots Y'_p]$. Soit s le rang des colonnes et $i_1, \dots, i_s$ les indices de pivot de M'.

Le sous-espace V de K^n engendré par les Y_i coïncide avec celui V' engendré par les Y'_i. Comme la dimension de V est égale à celle du rang des lignes et celle de V' au rang des colonnes, on en déduit que ces deux rangs coïncident.

Plus précisément, notons W_i, pour $0 \leqslant i \leqslant n$, le sous-espace vectoriel $\{0\}^i \times K^{n-i}$ de K^n défini par l'annulation des i premières coordonnées. Par définition, si $1 \leqslant j \leqslant s$, Y'_j appartient à W_{i_j}, et $Y'_j = 0$ si $j > s$. On constate ainsi que si $i_j \leqslant i < i_{j+1}$, alors $W_i \cap V$ est engendré par les vecteurs $Y'_1, \dots, Y'_j$. Par suite, $\dim(W_i \cap V) = j$ si $i_j \leqslant i < i_{j+1}$.