À propos des matrices échelonnées

Antoine Ducros
appendice au cours de *Géométrie affine et euclidienne* dispensé à l'Université Paris 6

Année universitaire 2011-2012

Introduction

Soit k un corps, soit $\mathscr{E}$ un k-espace affine de dimension finie, et soit $\mathscr{R}$ un repère cartésien de $\mathscr{E}$. Il y a, en pratique, deux façons de décrire un sous-espace affine $\mathscr{F}$ de $\mathscr{E}$ en coordonnées dans $\mathscr{R}$.

(i) *On peut en donner un système d'équations cartésiennes (affines), c'est-à-dire encore le décrire comme l'ensemble des antécédents d'un point par une application affine.* Par exemple, le système d'équations

$$\left\{\begin{array}{ccccccc} x & + & y & - & z & = & 1 \\ x & - & y & + & 3z & = & 3 \end{array}\right.$$

définit une droite de $\mathbb{R}^3$ (muni de son repère canonique) ; il la décrit très précisément comme l'ensemble des antécédents du point $(1,3)$ par l'application affine

$$\mathbb{R}^3 \to \mathbb{R}^2, (x,y) \mapsto (x+y-z, x-y+3z).$$

(ii) *On peut en donner un paramétrage (affine), c'est-à-dire encore le décrire comme l'image d'une application affine.* Par exemple,

$$\{(2-5s+t, 2+s-t, 4+2s-t)\}_{(s,t)\in\mathbb{R}^2}$$

définit un plan de $\mathbb{R}^3$ (muni de son repère canonique), et le décrit plus précisément comme l'image de l'application affine

$$\mathbb{R}^2 \to \mathbb{R}^3, (s,t) \mapsto (2-5s+t, 2+s-t, 4+2s-t).$$

Remarquons que se donner un paramétrage affine de $\mathscr{F}$ revient à en donner un point (ses coordonnées sont les termes constants du paramétrage) et une famille génératrice de l'espace directeur (formée par les ≪vecteurs de coefficients≫ de chacun des paramètres). Ainsi, dans l'exemple ci-dessus, le sous-espace affine considéré est égal à

$$\left\{ \underbrace{(2,2,4)}_{\text{termes constants}} + s\cdot \underbrace{(-5,1,2)}_{\text{coefficients de } s} + t\cdot \underbrace{(1,-1,-1)}_{\text{coefficients de } t} \right\}_{(s,t)\in\mathbb{R}^2}$$

et est donc le sous-espace affine de $\mathbb{R}^3$ passant par le point $(2,2,4)$ et dirigé par $\mathrm{Vect}((-5,1,2),(1,-1,-1))$.

Ces deux types de descriptions ont leurs mérites propres.

Une description par équations cartésiennes permet de savoir immédiatement si un point donné appartient à $\mathscr{F}$: il suffit de regarder si ses coordonnées satisfont les équations. Ainsi on voit immédiatement que le point $(1, 1, 1)$ appartient à la droite donnée en (i), mais pas le point $(2, 2, 2)$ (qui ne satisfait pas la seconde équation). Une description paramétrique ne permet pas de traiter ce genre de question aussi simplement ; par exemple, on ne peut pas dire *instantanément* si $(1, 5, -1)$ appartient ou non au plan décrit en (ii).

Une description paramétrique fournit, par sa forme même, une liste *explicite* de tous les points de $\mathscr{F}$. Si l'on connaît un système $\mathscr{S}$ d'équations cartésiennes de $\mathscr{F}$, une telle liste peut être vue comme une liste explicite de toutes les solutions de $\mathscr{S}$; on ne peut pas en exhiber une *instantanément* à partir de $\mathscr{S}$.

Le but de ce petit texte est de fournir une méthode systématique de passage d'une description à l'autre. Elle n'est pas compliquée – elle consiste simplement à utiliser le pivot de Gauß pour mettre, par manipulations élémentaires *sur les lignes*, une matrice donnée sous une forme particulière, dite *échelonnée (en lignes)*, mais se révèle redoutablement puissante ; expliquons succinctement en quoi.

• Elle marche *aussi bien* pour passer d'une description paramétrique aux équations cartésiennes que pour aller dans l'autre sens.

• Elle fournit dans chaque cas une description aussi simple que possible, dans le sens suivant.

Supposons tout d'abord que l'on parte d'un système *quelconque* d'équations cartésiennes $\mathscr{S}$ décrivant un sous-espace affine $\mathscr{F}$ de $\mathscr{E}$. L'échelonnement fournit alors un point et une *base* (et pas seulement une famille génératrice) de l'espace directeur de $\mathscr{F}$; en d'autres termes, il permet d'obtenir une description de l'ensemble des solutions de $\mathscr{S}$ avec un nombre minimal de paramètres.

Supposons maintenant que l'on parte d'une description paramétrique *quelconque* de $\mathscr{F}$. L'échelonnement fournit alors un système *minimal* d'équations cartésiennes de $\mathscr{F}$, c'est-à-dire que les équations linéaires qui leur sont associées forment une famille libre.

• Elle permet de *simplifier* la description d'un sous-espace affine par un système d'équations cartésiennes : si l'on se donne un système d'équations cartésiennes $\mathscr{S}$ d'un sous-espace affine $\mathscr{F}$ de $\mathscr{E}$, elle fournit à partir de $\mathscr{S}$ un système d'équations cartésiennes $\mathscr{S}'$ de $\mathscr{F}$ dont les équations linéaires associées forment une famille libre ; de plus, les équations qui constituent $\mathscr{S}'$ sont particulièrement simples.

• Elle permet d'inverser une matrice, et donc de calculer la réciproque d'une application affine.

• Mentionnons toutefois une limite de cette méthode : si l'on se donne un point et une famille génératrice de l'espace directeur de $\mathscr{F}$, on ne peut pas directement récupérer par échelonnement en lignes une base de l'espace directeur ; il faut ou bien procéder deux fois de suite à un tel échelonnement, ou bien utiliser (et une fois suffit alors), sa version transposée, c'est-à-dire l'échelonnement en colonnes.

1 Matrices échelonnées

À propos des matrices vides

Si n est un entier, on note $\{1, \ldots, n\}$ l'ensemble des entiers compris entre 1 et n ; il est donc vide dès que $n = 0$.

Si n et m sont deux entiers, une matrice de taille (n, m) à coefficients dans k est une application $(i, j) \mapsto a_{ij}$ de $\{1, \ldots, n\} \times \{1, \ldots, m\}$ dans k. Cette définition garde un sens lorsque n ou m est nul : l'ensemble $\{1, \ldots, n\} \times \{1, \ldots, m\}$ est alors vide, et il existe une et une seule application de l'ensemble vide dans k (et plus généralement dans n'importe quel ensemble E) : c'est l'inclusion, qu'on appelle aussi l'application vide.

Cela a donc un sens de parler de matrice de taille (m, n) si m ou n est nul ; il existe une et une seule matrice de cette taille, parfois appelée matrice vide.

Contrairement à ce que l'on pourrait croire au premier abord, inclure ces étranges bestioles dans la théorie n'a pas pour but de compliquer la vie, mais bel et bien de la simplifier. Cela assure par exemple que tous les raisonnements d'algèbre linéaire reposant sur des arguments matriciels restent valables même lorsqu'un des espaces en jeu est de dimension nulle, et évite donc des distinctions fastidieuses de cas particuliers le plus souvent triviaux.

De même, cela peut permettre, par exemple, au cours d'un raisonnement sur une matrice de taille (n, m) avec $n \geqslant 1$, de considérer la sous-matrice de taille $(n-1, m)$ obtenue en retirant la première ligne, sans avoir à distinguer le cas où n est égal à 1.

Nous avons donc choisi, dans ce texte, d'autoriser les matrices de taille (n, m) avec n ou m nul. Le lecteur qui n'aime pas ça peut ou bien ne pas y penser – c'est le plus raisonnable –, ou bien supposer que n et m sont strictement positifs, à charge pour lui d'initialiser certaines récurrences à $n = 1$, alors qu'elles démarrent ici à $n = 0$.

Matrices échelonnées

Soient n et m deux entiers, et soit $A \in \mathrm{M}_{n,m}(k)$. On dit que A est *échelonnée (selon les lignes)* si elle possède les propriétés suivantes :

• pour tout $i \in \{1, \ldots, n\}$, si $L_i = 0$ alors $L_j = 0$ pour tout $j > i$;

• pour tout $i \in \{1, \ldots, n\}$, si $L_i \neq 0$, son premier coefficient non nul est égal à 1, et est appelé un *pivot* de la matrice A ;

• pour tout $i \in \{1, \ldots, n\}$ tel que $L_i \neq 0$ et tout $j > i$ alors ou bien L_j est nulle, ou bien le pivot de L_j est situé *strictement à droite* de celui de L_i ;

• dans la colonne d'un pivot, tous les termes sont nuls hormis le pivot lui-même.

On définit de même la notion de matrice échelonnée selon les colonnes ; une matrice est échelonnée selon les lignes si et seulement si sa transposée est échelonnée selon les colonnes. Dans ce qui suit, une matrice échelonnée sera toujours, sauf mention expresse du contraire, échelonnée selon les lignes.

Par exemple,
$$\begin{pmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
est échelonnée ; ses pivots sont situés en $(1,1)$, $(2,3)$ et $(3,5)$.

Si $A = (a_{ij})$ est une matrice échelonnée, ses lignes non nulles forment une famille libre. En effet, soit I l'ensemble des indices des lignes non nulles de A, et supposons que $\sum_{\in\in I} \lambda_i L_i = 0$. Fixons i_0, et soit j le numéro de la colonne contenant le pivot de L_{i_0}. On a en particulier $\sum_{i\in I} \lambda_i a_{ij} = 0$. Mais comme tous les termes de la colonne j sont nuls, à part le pivot de L_{i_0} qui vaut 1, il vient $\lambda_{i_0} = 0$.

En conséquence, le rang d'une matrice échelonnée est égal au nombre de ses lignes non nulles, c'est-à-dire encore au nombre de ses pivots. Ainsi, la matrice ci-dessus est de rang 3.

Soit A une matrice carrée de taille n, supposée échelonnée. Elle est de rang n, c'est-à-dire inversible, si et seulement si elle a n pivots. Il est immédiat, compte-tenu de la définition même d'une matrice échelonnée, que cela se produit si et seulement si $A = I_n$.

Appelons *opération élémentaire sur les lignes de* A une opération de l'un des types suivants :

- $L_i \longleftrightarrow L_j$, avec $i \neq j$;
- $L_i \to \alpha L_i$ avec $\alpha \in k^*$;
- $L_i \to L_i + \alpha L_j$, avec $j \neq i$ et $\alpha \in k$.

Chacune de ces opérations consiste à multiplier A *à gauche* par une matrice inversible convenable : celle déduite de l'identité par l'opération considérée.

Lemme. *Soit $A \in M_{n,m}(k)$. On peut la transformer en une matrice échelonnée par une suite finie d'opérations élémentaires sur les lignes.*

Démonstration. La preuve que nous allons donner va essentiellement consister à décrire un algorithme permettant de réaliser la réduction souhaitée. On raisonne par récurrence sur le nombre m de colonnes.

Si $m = 0$ la matrice est vide, et donc échelonnée.

Supposons $m > 0$ et la propriété vraie en rang inférieur.

Premier cas. Supposons que la première colonne de A est nulle ; en vertu de l'hypothèse de récurrence, on peut échelonner la matrice $(a_{ij})_{1\leqslant i\leqslant n, 2\leqslant j\leqslant m}$ par opérations élémentaires sur les lignes. Lorsqu'on applique les opérations en question à A, on obtient une matrice dont la première colonne est nulle, et dont le bloc de droite de taille $(n, m-1)$ est échelonné ; cette matrice est échelonnée, et la preuve est terminée.

Second cas. Supposons qu'il existe i tel que $a_{i1} \neq 0$. Quitte à échanger L_i et L_1 si $i \neq 1$, on se ramène au cas où $a_{11} \neq 0$; puis, en multipliant L_1 par a_{11}^{-1}, au cas où $a_{11} = 1$.

Enfin, en remplaçant pour tout $i > 1$ la ligne L_i par $L_i - a_{i1}L_1$, on obtient une matrice telle que $a_{11} = 1$ et $a_{i1} = 0$ pour $i > 1$.

Soit B le bloc $(a_{ij})_{2\leqslant i\leqslant n, 2\leqslant j\leqslant m}$ de A. Par l'hypothèse de récurrence, on peut échelonner B par opérations élémentaires sur les lignes. En appliquant les opérations correspondantes à A, on obtient une matrice de la forme suivante :

- son terme en haut à gauche est égal à 1, et est le seul terme non nul de la première colonne ;
- son bloc $(n-1, m-1)$ en bas à droite est échelonné.

La matrice ainsi obtenue n'est pas forcément échelonnée : dans la colonne d'un pivot de B le terme situé sur la première ligne de A pourrait être non nul. On y remédie comme suit.

Soient $i_1, \ldots, i_r$ les numéros (croissants) des lignes de A comportant un pivot de B ; pour tout $\ell \in \{1, \ldots, r\}$ notons j_ℓ le numéro de la colonne de A contenant le pivot de B situé sur $L_{i\ell}$.

Faisons subir à A les opérations $L_1 \to L_1 - a_{1j_1}L_{i_1}, L_1 - a_{1j_2}L_{i_2}$, etc. L'opération $L_1 \to L_1 - a_{1j_\ell}L_{i_\ell}$ remplace a_{1j_ℓ} par 0, et ne modifie aucun des a_{1s} pour $s < j_\ell$ puisque $a_{i_\ell s} = 0$ pour tout $s < j_\ell$. Par conséquent, la matrice A obtenue au terme de ces transformations est bien échelonnée. $\square$

Le noyau d'une matrice échelonnée

Soit A une matrice de taille (n, m) échelonnée. Interprétons-la comme la matrice d'un endomorphisme de k^m dans k^n, munis de leurs bases canoniques respectives. Soit J l'ensemble des indices des colonnes comportant un pivot et soit J' son complémentaire.

Pour tout $j \in J$ on note $i(j)$ la ligne qui contient le pivot de la colonne j.

Par définition même d'une matrice échelonnée, le noyau E de A peut être décrit, en coordonnées dans la base canonique de k^m, par le système d'équations cartésiennes

$$x_j + \sum_{\ell > j, \ell \in J'} a_{i(j)\ell}x_\ell = 0, \;\; j \in J,$$

soit encore

$$(*) \quad x_j = -\sum_{\ell > j, \ell \in J'} a_{i(j)\ell}x_\ell, \;\; j \in J.$$

Reconstitution de la matrice A à partir de E. Nous allons expliquer comment A peut être retrouvée en termes des propriétés du sous-espace vectoriel E de k^m ; il en résultera que *deux matrices échelonnées qui ont même noyau coïncident.*

Si $m = 0$ alors $E = k^m = 0$, la matrice A est nécessairement vide, et il n'y a rien à faire. On suppose à partir de maintenant que $m > 0$.

Reconstitution de l'ensemble J. On va en fait reconstituer J' – cela revient au même puisque J est le complémentaire de J', mais J' est plus facile à décrire directement. Si S est une partie de $\{1, \ldots, m\}$ on dira que S est *libre relativement à E* si pour toute famille $(a_j)_{j\in S}$ d'éléments de k il existe un élément $(x_j)_{1\leqslant j\leqslant m}$ de E tel que $x_j = a_j$ pour tout $j \in S$.

En vertu de $(*)$, l'ensemble J' peut alors être décrit comme suit, par un procédé récursif descendant :

• l'entier m appartient à J' si et seulement si $\{m\}$ est libre relativement à E ;
• pour tout $j \in \{1, \ldots, m-1\}$ on a $j \in J'$ si et seulement si l'ensemble $\{j\} \cup (\{j+1, \ldots, m\} \cap J')$ est libre relativement à E.

Reconstitution des indices $i(j)$. Il découle de la définition d'une matrice échelonnée que le ℓ-ième élément de J (pour l'ordre usuel sur les entiers) est nécessairement situé sur la ligne ℓ, ce qui montre que la fonction $j \mapsto i(j)$ est uniquement déterminée une fois que l'on connaît J.

Fin de la reconstitution de A. Il reste à déterminer les $a_{i(j)\ell}$ pour $j \in J$ et $\ell \in J' \cap \{j+1, \ldots, m\}$. Fixons donc un tel couple (j, ℓ). Pour tout $a \in K$, soit e_a l'élément de k^m dont la j-ième coordonnée est égale à a, dont la ℓ-ième coordonnée est égale à -1, et dont les autres coordonnées sont nulles. Il résulte de $(*)$ que $a_{i(j)\ell}$ est l'unique élément a de k tel que $e_a \in E$, ce qui termine la reconstitution.

Corollaire : unicité de la forme échelonnée

Soient n et m deux entiers et soit $A \in M_{n,m}(k)$. On a vu plus haut que l'on peut, par une suite finie d'opérations élémentaires sur ses lignes, transformer A en une matrice échelonnée B. Compte-tenu de l'interprétation matricielle d'une telle opération, cela implique l'existence d'une matrice inversible P de taille n telle que $PA = B$. Si X est un vecteur colonne de taille m on a

$$AX = 0 \iff PAX = 0,$$

puisque P est inversible. Autrement dit, le noyau de PA, c'est -à-dire le noyau de B, est égal à celui de A. Comme une matrice échelonnée peut être reconstruite à partir de son noyau par le procédé décrit plus haut, il en découle *l'unicité de* B : si l'on échelonne A par opérations élémentaires sur ses lignes, la matrice obtenue ne dépend pas de la suite d'opérations effectuées.

2 Utilisation de l'échelonnement

Inversion de matrice

Soit n un entier et soit A une matrice carrée de taille n. On peut la transformer en une matrice échelonnée par une suite d'opérations élémentaires, chacune d'elle consistant à multiplier A à gauche par une matrice inversible d'un type particulier. On peut ainsi écrire $P_r P_{r-1} \ldots P_1 A = B$, où P_i est la matrice correspondant à la i-ième opération élémentaire que l'on a effectuée, et où B est échelonnée.

Comme les P_i sont inversibles, le rang de A est égal à celui de B, dont on a vu qu'il est lui-même égal au nombre de lignes non nulles de B. En particulier, A est inversible si et seulement si B est inversible. On a vu plus haut qu'une matrice carrée échelonnée est inversible si et seulement si elle est égale à l'identité ; par conséquent, A est inversible si et seulement si B est égale à l'identité. Dans ce cas, l'inverse de A est égale à $P_r P_{r-1} \ldots P_1$; compte-tenu de l'interprétation des P_i comme matrices d'opérations élémentaires sur les lignes, on voit que A^{-1}

s'obtient en faisant subir à la matrice I_n les transformations que l'on a fait subir à A.

Récapitulation. Pour déterminer si une matrice A est inversible et, le cas échéant, calculer son inverse, on procède comme suit :

• on échelonne A par opération élémentaires sur les lignes ; on fait subir à la matrice I_n la même suite d'opérations élémentaires ; on note B (resp. C) la matrice ainsi obtenue à partir de A (resp. I_n) ;

• si $B = I_n$, la matrice A est inversible et $A^{-1} = C$;

• si $B \neq I_n$ alors A n'est pas inversible.

Échelonnement, image et antécédents

Soient n et m deux entiers et soit A une matrice appartenant à $M_{n,m}(k)$. Fixons un vecteur colonne B de longueur n, dont on note $b_1, \ldots, b_n$ les composantes.

Le but de ce qui suit est d'utiliser l'échelonnement pour donner : une équation cartésienne de l'image de $F\ X \mapsto AX + B$; et pour tout point de cette image, une description paramétrique de l'ensemble de ses antécédents.

Pour cela, on procède comme suit. Soit Y un vecteur colonne de taille n de coordonnées $y_1, \ldots, y_n$. On échelonne la matrice A en lui faisant subir une suite d'opérations sur les lignes ; si $P_1, \ldots, P_r$ désignent les matrices codant chacune de ces opérations, on a $P_r P_{r-1} \ldots P_1 A = C$, où C est échelonnée.

Posons $Y' = P_r P_{r-1} ... P_1 Y$ et $B' = P_r P_{r-1} ... P_1 B$. Le vecteur colonne Y' (resp. B') est obtenu en faisant subir à Y (resp. B) les transformations infligées à A ; le i-ème terme de Y' est égal à $\ell_i(y_1, \ldots, y_n)$ pour une certaine forme linéaire ℓ_i ; notons b'_i le i-ième terme de B'.

Soit X un vecteur colonne de longueur m de coordonnées $x_1, \ldots, x_m$. On a

$$AX + B = Y \iff P_r P_{r-1} \ldots P_1 AX + P_r P_{r-1} \ldots P_1 B = P_r P_{r-1} \ldots P_1 Y,$$

soit encore $CX + B' = Y'$.

Soit r le nombre de lignes non nulles de C ; les lignes non nulles de C sont alors exactement les lignes dont le numéro appartient à $\{1, \ldots, r\}$. Soit J l'ensemble des numéros des colonnes des pivots. Pour tout $i \in \{1, \ldots, r\}$, on note $\pi(i)$ le numéro de la colonne du pivot de la i-ième ligne ; on a donc $J = \{\pi(i)\}_{1 \leqslant i \leqslant r}$.

L'égalité $AX + B = Y$ équivaut par ce qui précède à $CX + B' = Y'$, c'est-à-dire au système d'équations

$$(**) \left\{ \begin{array}{lcll} x_{\pi(i)} & = & -\left(\displaystyle\sum_{j > \pi(i), j \notin J} c_{ij} x_j \right) + \ell_i(y_1, \ldots, y_n) - b'_i & (1 \leqslant i \leqslant r) \\ 0 & = & \ell_i(y_1, \ldots, y_n) - b'_i & (r < i \leqslant n) \end{array} \right. .$$

Un système d'équations cartésiennes de F. Nous allons montrer que Y appartient à F si et seulement si $\ell_i(y_1, \ldots, y_n) = b'_i$ pour tout $i \in \{r+1, \ldots, n\}$. Le vecteur Y appartient à F si et seulement si le système $AX + B = Y$ a une solution en X, donc si et seulement si $CX + B' = Y'$ a une solution en X.

Supposons que ce soit le cas ; on déduit alors de $(**)$ que $\ell_i(y_1,\ldots,y_n)=b'_i$ pour tout $i\in\{r+1,\ldots,n\}$. Réciproquement, supposons que $\ell_i(y_1,\ldots,y_n)=b'_i$ pour tout $i\in\{r+1,\ldots,n\}$, et soit X le vecteur colonnes de coordonnées $x_1,\ldots,x_m$ définies comme suit :

- $x_j=0$ pour tout $j\in\{1,\ldots,m\}-J$;
- $x_{\pi(i)}=\ell_i(y_1,\ldots,y_n)-b'_i$ pour tout $i\in\{1,\ldots,r\}$.

Le système $(**)$ est alors satisfait, ce qui signifie exactement que $CX+B'$ est égal à Y' ; ainsi, Y appartient à F.

Par conséquent, $\{l_i(y_1,\ldots,y_n)=b'_i\}_{r+1\leqslant i\leqslant n}$ est un système d'équations cartésiennes de F. Le rang de A est égal au nombre de lignes non nulles de C, c'est-à-dire à r ; la dimension de F est donc r.

Le système $\{l_i(y_1,\ldots,y_n)=b'_i\}_{r+1\leqslant i\leqslant n}$ est un système de $n-r$ équations affines qui décrit un sous-espace affine de dimension r de k^n (identifié à l'ensemble des vecteurs colonnes de longueur n). Il s'ensuit que les parties linéaires de ces équations sont linéairement indépendantes : on a ainsi obtenu un système d'équations cartésiennes de F qui est minimal.

Description des antécédents d'un point de F. Soit $Y\in F$; cela signifie d'après ce qui précède que $\ell_i(y_1,\ldots,y_n)=b'_i$ pour tout $i\in\{r+1,\ldots,n\}$. Le système $(**)$ fournit alors par sa forme même une description paramétrique de l'ensemble E des antécédents de Y, c'est-à-dire de l'ensemble des X tels que $AX+B=Y$ ou, ce qui revient au même, tels que $CX+B'=Y'$: les paramètres sont les x_i pour $i\notin J$; ceux-ci étant fixés, on a pour tout $i\in\{1,\ldots,r\}$ l'égalité
$x_{\pi(i)}=-\left(\sum_{j>\pi(i),j\notin J}c_{ij}x_j\right)+\ell_i(y_1,\ldots,y_n)-b'_i.$

On voit que l'on obtient une description au moyen de $m-r$ paramètres (le cardinal de J étant égal à celui de I, et donc à r) ; en d'autres termes, on obtient par ce biais un point de E et une famille génératrice de cardinal r de son espace directeur. Par ailleurs, l'espace directeur en question s'identifie au noyau de A, et est donc de dimension $m-r$, puisque A est de rang r. La famille génératrice en question obtenue est en conséquence une base ; autrement dit, on a exhibé une description paramétrique de E avec un nombre minimal de paramètres.

Un point de vue légèrement différent : simplification d'un système d'équations cartésiennes de E. On peut donner une autre interprétation de la description de E ainsi obtenue : *a priori*, l'espace E est défini comme l'ensemble des antécédents de Y, donc comme l'ensemble des solutions (en X) de l'équation $AX+B=Y$; il est donc décrit par un système de n équations affines. Le système $AX+B=Y$ est équivalent à $CX+B'=Y'$. Dans celui-ci, d'après ce qui précède, les lignes numérotées de $r+1$ à n sont des égalités entre scalaires (elles ne font plus intervenir les variables x_j). Ne subsistent donc que les r premières lignes, dont les parties linéaires forment une matrice échelonnée ; elles sont donc linéairement indépendantes. On a ainsi obtenu une nouvelle description de E par un système d'équations cartésiennes qui a un nombre minimal d'équations et est de surcroît particulièrement simple.

Échelonnement en colonnes et extraction d'une base

Soient m et n deux entiers, soit A une matrice appartenant à $M_{nm}(k)$ et soit B un vecteur colonne de longueur n. Soit F l'image de $X \mapsto AX + B$. Décrire F comme l'image de $X \mapsto AX + B$ revient à en donner une description paramétrique, à m paramètres (les coordonnées de l'espace source) ; ou, si l'on préfère, à en donner un point (à savoir B) et une famille génératrice de l'espace directeur (les colonnes de A).

Nous allons expliquer dans ce qui suit comment donner une description paramétrique de F avec un nombre minimal de paramètres – où, ce qui revient au même, comment en donner une *base* de l'espace directeur.

L'échelonnement en lignes préserve le noyau d'une matrice (*cf. supra*), mais n'a aucune raison de préserver son image ; il n'est donc pas adapté au problème posé.

Pour résoudre celui-ci, nous allons échelonner A en colonnes, par manipulations élémentaires sur ses colonnes ; pour s'assurer que c'est possible on peut ou bien reprendre la preuve de l'existence d'un échelonnement en ligne, et échanger les termes «ligne» et «colonne» dans la démonstration ; ou bien utiliser l'existence de l'échelonnement en lignes en l'appliquant à ${}^tA$.

Toute manipulation élémentaire sur les colonnes de A revient à la multiplier à droite par une matrice inversible convenable (obtenue en appliquant la manipulation en question à l'identité). Il existe donc une matrice inversible P et une matrice échelonnée en colonnes C telles que $C = AP$.

Un vecteur Y appartient à F si et seulement si il existe un vecteur colonne X de longueur n tel que $AX + B = Y$. Ceci est équivalent à l'existence d'un vecteur X' tel que $APX' + B = Y$: en effet, si un tel X' existe, on peut prendre $X = PX'$; et si un tel X existe, on peut prendre $X' = P^{-1}X$. Par conséquent, F est l'image de l'application $X \mapsto CX + B$. Comme C est échelonnée en colonne, ses colonnes non nulles sont linéairement indépendantes ; elles constituent donc une base de l'espace directeur de F.

Alternativement, la description de F comme l'image de $X \mapsto CX + B$ en fournit un paramétrage, dans lequel les paramètres qui interviennent effectivement sont les coordonnées de l'espace source correspondant aux colonnes non nulles de C ; si r désigne le rang de A, il y a r telles colonnes, et on a ainsi obtenu une description de F au moyen du nombre minimal de paramètres possible, à savoir r.