Université Paris 7 – Denis-Diderot
Support de cours – 04/02/2006 – M302 Géométrie

1) **Qu'est-ce qu'une matrice échelonnée** – La définition d'une matrice échelonnée est en apparence un peu compliquée, mais est au fond bien simple.

Une matrice échelonnée de $\mathbb{K}^{p\times n}$, c'est-à-dire ayant p lignes et n colonnes, est une matrice soumises à quatre conditions :
a) si une ligne est nulle, toutes celles qui sont en dessous sont nulles également (le nombre r de lignes non nulles est appelé le rang de la matrice échelonnée) ;
b) le premier terme non nul dans une ligne non nulle est obligatoirement égal à un (un tel terme est appelé marquer d'échelon, et il y a donc autant de marqueurs d'échelon qu'il y de lignes non nulles, c'est-à-dire r, et pour nommer un marqueur d'échelon on désignera sa ligne et pour le situer on a besoin de préciser sa colonne) ;
c) les termes qui sont au dessus d'un marqueur d'échelon (au delà du premier) sont nuls ;
d) les numéros de colonnes des échelons successifs vont en croissant strictement, autrement dit, ils occupent dans la matrice échelonnée une position en escalier descendant le long de la direction NO-SO.

Il s'ensuit qu'un marqueur d'échelon est le seul terme non nul de sa colonne.

Exercice – Trouver les trois matrices échelonnées cachées parmi le groupe suivant de matrices :

$$M = \begin{bmatrix} 1 & 3 & 2 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix},\ M = \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},\ M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},\ M = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$M = \begin{bmatrix} 1 & -3 & -2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},\ M = \begin{bmatrix} 1 & -3 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 22 & -15 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ et } M = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2) **Considérations sur la dualité** – L'algèbre linéaire commença quand Gauss, s'ennuyant, fit quelques *manipulations élémentaires* sur les équations d'un système (linéaire). On oublie cela de nos jours, car les étudiants apprennent très tôt ce qu'est un espace vectoriel, comment on additionne des vecteurs, que cela rappelle la composition des forces en physique. Mais la naissance de l'idée d'espace vectoriel provient justement du fait que les équations (ou tout au moins les membres de gauche d'icelles) sont aussi des vecteurs ! La prise de conscience que les manipulations sur

les vecteurs de $\mathbb{R}^n$ ou sur les formes linéaires sur $\mathbb{R}^n$ sont de même nature a fait inventer la *structure* d'espace vectoriel.

La contribution de Gauss dans l'étude des systèmes linéaires est en quelque sorte le terreau sur lequel a germé l'idée de structure d'espace vectoriel. Le problème du "contrôle" des sous-espaces vectoriels de $\mathbb{R}^n$ revient alors dans ce contexte à saisir dans un premier temps comment parler d'un sous-espace au moyen d'un système d'équations ou d'un système de générateurs (et l'on parlera alors d'un contrôle par équations ou d'un contrôle par générateurs) et, dans un second temps, comment passer d'un type de contrôle à un autre. En langage moderne, il s'agit d'exhiber une base du sous-espace F et une base de l'orthogonal $F^\perp$ de F dans le dual, et de disposer d'un procédé qui permette de passer d'une système définissant (par équations ou par générateurs) l'un de ces sous-espaces en une base de l'un et de l'autre. On dira que l'on a *affiné le contrôle* si partant d'un système (en général) redondant (d'équations ou de générateurs) on produit un système minimal (de même nature), par remplacement ou par extraction. Et l'on dit que l'on a *inversé le contrôle* si partant d'un système d'équations de F on exhibe une base de F, ou, si partant d'un système générateur de F, on exhibe un système d'équatios minimal de F, c'est-à-dire une base de $F^\perp$. Le langage qui vient d'être utilisé met en évidence la symétrie de nature qui unit les vecteurs de l'espace vectoriel E et les formes linéaires sur E, qui, dans la pratique, sont matérialisées par les équations. Ce langage s'applique pareillement aux vecteurs et aux équations, ou plus précisément à leurs premiers membres. Il met en évidence, répétons-le, l'idée d'espace vectoriel et au delà de celle-ci, l'idée de structure, qui unit des objets de nature visiblement disparate, mais cependant si semblables par la parenté des liens qui s'y opèrent.

Expliquons dans ce qui suit pourquoi la donnée d'un sous-espace vectoriel F de $\mathbb{K}^n$ par un système de p équations en n variables signifie que l'on a défini F en définissant son orthogonal $F^\perp$ et ce au moyen de p générateurs. Plus précisément, se donner F comme l'ensemble des solutions d'un système de p équations à n inconnues, c'est définir F comme intersection de p hyperplans $H_1, \ldots, H_p$ de $\mathbb{K}^n$, chacun des hyperplans étant le noyau de la forme linéaire correspondante, qui apparaît à la p-ième ligne de notre système d'équations :

$$F \;=\; \operatorname{Ker} \varphi_1 \cap \operatorname{Ker} \varphi_2 \cap \cdots \cap \operatorname{Ker} \varphi_p \,.$$

Mais

$$F^\perp = \mathbb{K}\varphi_1 + \mathbb{K}\varphi_2 + \cdots + \mathbb{K}\varphi_p = \langle \varphi_1; \varphi_2; \ldots; \varphi_p \rangle,$$

car l'orthogonal d'une intersection est, en dimension finie, égal à la somme des orthogonaux. Autrement dit, le sous-espace $F^\perp \subseteq E^\star$ est engendré par la famille des formes (φ_i).

Notons en passant que lorsque F est défini par p équations

$$\varphi_1(x_1, \ldots, x_n) = 0, \quad \varphi_2(x_1, \ldots, x_n) = 0, \; \ldots, \quad \varphi_p(x_1, \ldots, x_n) = 0,$$

toute forme linéaire qui s'annule sur F est combinaison linéaire des formes $\varphi_1, \dots, \varphi_i$.

3) Les quatre problèmes – Cela dit, affiner le contrôle sur F revient à proposer un système de $r = \dim F^{\perp} = \dim E - \dim F$ équations pour définir F. Dans la pratique, la matrice A du sytème d'équations est échelonnée en $\acute{E}L(A)$, qui est de rang r égal au nombre de ses lignes non nulles. Comme $F = \operatorname{Ker} A = \operatorname{Ker} \acute{E}L(A)$, on remplace le système de départ par le système d'équations minimal fourni par les formes linéaires non nulles lues dans l'échelonnée.

Le pivot de Gauss fournit, à partir de la matrice A que jouxte à sa gauche la matrice I_p, fournit donc non seulement $\acute{E}L(A)$ mais aussi une matrice $P \in GL(p, \mathbb{K})$ telle que

$$PA = \acute{E}L(A)\,.$$

Les lignes de P qui font face aux $p-r$ lignes nulles de $\acute{E}L(A)$ une fois multipliées par A donnent 0. Ce sont donc des formes linéaires qui s'annulent sur les colonnes de A, donc sur l'image de A. Ces lignes extraites de la matrice inversible P sont donc aussi libres. Écrivons la suite exacte courte associée à l'endomorphisme $A : K^n \to \mathbb{K}^p$:

$$\operatorname{Ker} A \hookrightarrow \mathbb{K}^n \twoheadrightarrow \operatorname{Im}(A) \subseteq \mathbb{K}^p.$$

Étant en nombre exactement égal à la codimension de $\operatorname{Im}(A)$ dans K^p, ces formes linéaires en question fournissent donc une base de $\operatorname{Im}(A)^{\perp}$, c'est-à-dire un système d'équations minimal de l'image. Nous venons d'inverser le contrôle sur l'image, car l'image est fournie au départ par le système de générateurs que constituent les colonnes de la matrice A.

À cette étape nous pouvons procéder à un gain de plus : affiner le contrôle sur l'image, et cette fois en *extrayant* une base du système de générateurs fourni par les colonnes de A, plutôt que d'en proposer un en remplacement. La matrice échelonnée a les mêmes relations entre ses colonnes que la matrice A de départ : *une relation entre les colonnes* n'est-elle pas après tout et tout simplement une famille $(\lambda_1, \dots, \lambda_n)$ de $\mathbb{K}^n$ telle que $\sum_i \lambda_i C_i(A) = 0$? Comme $C_i(A) = A(e_i)$, où e_i est le i-ème vecteur de la base canonique de $\mathbb{K}^n$, on a donc que

$$0 = \sum_i \lambda_i C_i(A) = \sum_i \lambda_i A(e_i) = A(\sum_i \lambda_i e_i)\,.$$

Les colonnes de A qui occupent les mêmes positions que les marqueurs d'échelon dans la matrice $\acute{E}L(A)$ sont, d'une part en nombre r, qui est le bon nombre, et d'autre part, engendrent les autres colonnes : elles forment donc une base de $\operatorname{Im}(A)$.

Il reste à inverser le contrôle sur le noyau, et c'est cela qui intéresse au fond Gauss. Autrement dit, il s'agit de résoudre le système associé à A (ou disons celui auquel

A est associée). Cela veut dire trouver une famille de $n-r$ vecteurs $V_1, \ldots, V_{n-r}$ telle que toute solution du système soit une combinaison de ces vecteurs :

$$V \in \operatorname{Ker} A \quad \Longleftrightarrow \quad \exists\ \mu_1, \ldots, \mu_{n-r}, \ \mid\ V = \sum_{i=1}^{i=n-r} \mu_i V_i\,.$$

C'est ce qu'antan on appelait les équations paramétriques du sous-espace $\operatorname{Ker} A$.

C'est à cette seule étape que l'on s'autorise à manipuler les colonnes, et seulement celles de $\acute{E}L(A)$ (que l'on coiffe de la matrice I_n). On produit alors une matrice inversible $Q \in GL(n, K)$ telle que

$$\acute{E}L(A)\, Q = I_r(p, n)\,.$$

Les colonnes de cette matrice Q qui surplombent en fin de processus les colonnes nulles de $I_r(p,n)$, et qui sont en nombre $n-r$, sont d'une part indépendantes (car extraites de la matrice inversible Q), et sont, d'autre part, annulées par $\acute{E}L(A)$: elles forment donc une base du noyau d'icelle et donc de celui de A, qui est de dimension $n-r$. On a ainsi inversé le contrôle sur le noyau de A.

On aura noté que l'algorithme du pivot mené de son début à la fin produit de façon effective deux matrices inversibles $P \in GL(p, \mathbb{K})$ et $Q \in GL(n, \mathbb{K})$ telles que

$$PAQ = I_r(p, n)\,.$$

C'est "la" décomposition de Steinitz de $A \in \mathbb{K}^{p \times n}$.

Remarques
a) Le contrôle double d'un sous-espace vectoriel $F \in \mathbb{K}^n$, par des équations cartésiennes ou par des équations paramétriques, offrent deux manières assez dissimilaires pour penser celui-ci. Les équations cartésiennes fournissent un *test d'appartenance* à F alors que les équations paramétriques fournissent un moyen d'en *enfanter* autant d'éléments qu'il en est permis.
b) Retenons que lorsque l'on manipule (élémentairement), en compagnie de Gauss, les lignes d'un système linéaire, nous n'affectons heureusement ni ses solutions, ni les relations qui existent entre les colonnes de sa matrice A. Une solution ou une relation sont deux termes qui désignent enfin de compte une et même chose : un élément courant de $\operatorname{Ker} A$.
d) Notons bien que les matrices P et Q rencontrées plus haut ne sont pas uniques mais que $\acute{E}L(A)$ l'est tout comme, bien sûr, $I_r(p,n)$. Le secret de l'unicité de $\acute{E}L(A)$ réside dans le fait qu'on y encode les relations entre les colonnes de A.
c) Retenons aussi qu'un noyau est contrôlé au départ par un système d'équations, alors qu'une image est contrôlée au départ par un système de générateurs. Et les *quatre problèmes* avec lesquels on amuse les étudiants en L1, premier semestre, consistent donc à affiner et à inverser le contrôle sur l'un et l'autre de ces sous-espaces.