2.14 Forme normale de Smith et facteurs invariants (122, 142, 154, 162) [3]

Ce développement, qui m'a l'air très populaire dans la promo agreg 2024 de l'ENS (et c'est très bien !), généralise au cas des matrices à coefficients dans un anneau principal le fait suivant :

Théorème 2.35. Soit k un corps et $A \in \mathscr{M}_{m,n}(k)$. Alors, il existe un unique entier r et il existe $(P,Q) \in GL_m(k) \times GL_n(k)$ tel que :

$$P^{-1}AQ = \begin{pmatrix} I_r & \mathbf{O}_{r,q-r} \\ \mathbf{O}_{p-r,r} & \mathbf{O}_{p-r,q-r\cdot} \end{pmatrix}$$

Cet entier r est le rang de la matrice A et caractérise les orbites de l'action par équivalence : deux matrices A et B sont équivalentes si et seulement si elles ont le même rang.

Ce théorème se prouve algorithmiquement via le pivot de Gauss. On se base donc sur le fait que les éléments non-nuls de k sont inversibles pour diviser sans vergogne. Cependant, dans un anneau principal, on ne peut plus effectuer ces divisions. On effectue donc des substituts : si notre anneau est euclidien, on peut effectuer des divisions euclidiennes et donc remplacer l'algorithme du pivot de Gauss par un pivot de Gauss « avec restes » ou, si l'anneau est seulement principal, remplacer la division par l'application d'une relation de Bézout pour remplacer le coefficient que l'on voudrait éliminer par un PGCD. On a alors le résultat suivant :

Théorème 2.36 (Forme normale de Smith)**.** Soit A un anneau commutatif unitaire intègre que l'on suppose principal. Soient $m,n \in \mathbb{N}^*$ et soit $M \in \mathscr{M}_{m,n}(A)$. Alors il existe un entier s et des éléments $d_1,\dots,d_s \in A$ vérifiant les conditions de divisibilité :

$$d_1 \mid \dots \mid d_s$$

et des matrices $(P,Q) \in SL_m(A) \times SL_n(A)$ tels que :

$$P^{-1}MQ = \begin{pmatrix} d_1 & 0 & \cdots & 0 & \\ 0 & d_2 & \cdots & 0 & \mathbf{O}_{s,n-s} \\ \vdots & \ddots & \ddots & \vdots & \\ 0 & \cdots & 0 & d_s & \\ & & & & \\ & \mathbf{O}_{m-s,s} & & & \mathbf{O}_{m-s,n-s} \\ & & & & \end{pmatrix}.$$

De plus, le couple $(s,(d_1,\dots,d_s))$ est unique au sens suivant : s'il existe $(t,(d'_1,\dots,d'_t)) \in \mathbb{N}^* \times A^t$ vérifiant les mêmes hypothèses de divisibilité et s'il existe $(P',Q') \in GL_m(A) \times GL_n(A)$ tel que :

$$P'^{-1}MQ' = \begin{pmatrix} d'_1 & 0 & \cdots & 0 & \\ 0 & d'_2 & \cdots & 0 & \mathbf{O}_{t,n-t} \\ \vdots & \ddots & \ddots & \vdots & \\ 0 & \cdots & 0 & d'_t & \\ & & & & \\ & \mathbf{O}_{m-t,t} & & & \mathbf{O}_{m-t,n-t} \\ & & & & \end{pmatrix},$$

alors $s = t$ et pour tout $i \in [\![1,s]\!]$, d_i et d'_i sont associés. La matrice diagonale de ce théorème est appelée *forme normale de Smith* de la matrice A et les coefficients d_i sont appelés *facteurs invariants* de la matrice A.

Un corollaire immédiat est que deux matrices M et N de $\mathscr{M}_{m,n}(A)$ sont équivalentes si et seulement si elles ont les mêmes facteurs invariants. Ce théorème a des applications à la pelle ! Théorème de la base adaptée et structure des A-modules de type fini, résolution de systèmes d'équations diophantiennes linéaires, caractérisation des classes de similitude dans $\mathscr{M}_n(k)$ en calculant les facteurs invariants de $XI_n - A \in \mathscr{M}_n(k[X])$,...

Démonstration dans le cadre euclidien. Soit A un anneau euclidien, muni d'un stathme ν, c'est-à-dire une application :

$$\nu \quad : \quad A \setminus \{0\} \quad \longrightarrow \quad \mathbb{N}$$

telle que :

—

$$\forall (a,b) \in A \times A \setminus \{0\},\ \exists (q,r) \in A^2, \quad \left\{ \begin{array}{rcl} a & = & bq + r, \\ r = 0 & \text{ou} & \nu(r) < \nu(b). \end{array} \right.$$

— $\forall a, b \in A \setminus \{0\}, \quad \nu(ab) \geq \nu(a)$.

On a alors le fait élémentaire suivant : si $a, b \in A \setminus \{0\}$ sont tels que $a \mid b$ alors a et b sont associés si et seulement si $\nu(a) = \nu(b)$. De fait, on a donc que $a \in A$ est inversible si et seulement si $a \neq 0$ et $\nu(a) = \nu(1)$. On peut alors normaliser le stathme ν de sorte à ce que $a \in A^\times \iff \nu(a) = 1$.

Préparation : Dans la preuve du théorème, on sera amené à permuter les lignes ou les colonnes, mais le théorème que j'ai cité impose les matrices P et Q à être de déterminant 1. On va montrer le fait suivant :

Lemme 2.37 (Permuter deux lignes au signe près est une opération de déterminant 1)**.** Soit $M \in \mathscr{M}_{m,n}(A)$, de lignes notées $L_i, i \in [\![1, m]\!]$. Alors, pour tout $(i,j) \in [\![1,m]\!]^2$, la suite d'opérations élémentaires :

1. $L_i \longleftarrow L_i + L_j$,
2. $L_j \longleftarrow L_j - L_i$,
3. $L_i \longleftarrow L_i + L_j$

transforme la matrice M en la matrice :

$$M' = \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_j \\ L_{i+1} \\ \vdots \\ L_{j-1} \\ -L_i \\ L_{j+1} \\ \vdots \\ L_m \end{pmatrix}.$$

Démonstration. C'est vraiment pas compliqué ! Après l'étape 1 la ligne i de la matrice M est transformée en $L_i + L_j$. Ce sera notre "nouveau" L_i. Ainsi, en étape 2, la ligne j, qui n'a pas bougé en étape 1 est transformée en $L_j - (L_i + L_j) = -L_i$. Enfin, en étape 3, la ligne i qui est restée $L_i + L_j$ est transformée en $L_i + L_j - L_i = L_j$. Cela conclut ! □

Existence : Pour l'existence, on effectue une récurrence dans une récurrence. Plus précisément, on démontre le résultat par récurrence sur la taille de la matrice $t(M) := \max(m,n)$. Si $t(M) = 1$, alors $m = n = 1$, la matrice est diagonale et le résultat est bien vrai. Si $t(M) > 1$, alors on démontre le résultat à $t(M)$ fixé par récurrence sur la

quantité :

$$\nu(M) := \min_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} \nu(m_{i,j})$$

où $m_{i,j}$ désigne le coefficient en ligne i et colonne j de M. Si $\nu(M) = 1$, alors il existe un coefficient $m_{i,j}$ inversible. Ce coefficient va jouer notre rôle de pivot. Ainsi, en appliquant le lemme 2.37, il existe $(P_0, Q_0) \in SL_m(A) \times SL_n(A)$ tel que le coefficient en ligne 1 et colonne 1 de la matrice $P_0^{-1}MQ_0$ soit inversible. On peut donc considérer que $m_{1,1}$ est inversible. Ainsi, comme pour le pivot de Gauss, en effectuant les opérations :

$$\forall i \in [\![2, m]\!], \quad L_i \longleftarrow L_i - m_{1,1}^{-1} m_{i,1} L_1$$

puis :

$$\forall j \in [\![2, n]\!], \quad C_j \longleftarrow C_j - m_{1,1}^{-1} m_{1,j} C_1,$$

on obtient l'existence de matrices $(P_1, Q_1) \in SL_m(A) \times SL_n(A)$ telles que :

$$P_1^{-1}MQ_1 = \begin{pmatrix} m_{1,1} & \mathbf{O}_{1,n-1} \\ \mathbf{O}_{m-1,1} & M' \end{pmatrix}$$

avec $t(M') < t(M)$. Ainsi, en appliquant l'hypothèse de récurrence sur M', on a existence de matrices $(P', Q') \in SL_{m-1}(A) \times SL_{n-1}(A)$ et d'éléments $d_2 \mid \ldots \mid d_s \in A$ tels que :

$$P'^{-1}M'Q' = \begin{pmatrix} d_2 & 0 & \cdots & 0 & & \\ 0 & d_3 & \cdots & 0 & & \mathbf{O}_{s,n-1-s} \\ \vdots & \ddots & \ddots & \vdots & & \\ 0 & \cdots & 0 & d_s & & \\ & & & & & \\ & \mathbf{O}_{m-1-s,s} & & & & \mathbf{O}_{m-1-s,n-1-s} \\ & & & & & \end{pmatrix}.$$

En notant alors :

$$P_2 = \begin{pmatrix} 1 & \mathbf{O}_{1,m-1} \\ \mathbf{O}_{m-1,1} & P' \end{pmatrix} \in SL_m(A),$$

$$Q_2 = \begin{pmatrix} 1 & \mathbf{O}_{1,n-1} \\ \mathbf{O}_{n-1,1} & Q' \end{pmatrix} \in SL_n(A),$$

et $d_1 := m_{1,1}$, on a :

$$P_2^{-1}P_1^{-1}MQ_1Q_2 = \begin{pmatrix} d_1 & 0 & \cdots & 0 & & \\ 0 & d_2 & \cdots & 0 & & \mathbf{O}_{s,n-s} \\ \vdots & \ddots & \ddots & \vdots & & \\ 0 & \cdots & 0 & d_s & & \\ & & & & & \\ & \mathbf{O}_{m-s,s} & & & & \mathbf{O}_{m-s,n-s} \\ & & & & & \end{pmatrix}.$$

et on a bien la condition de divisibilité car $d_1 = m_{1,1}$ est inversible, et cela conclut donc l'initialisation. Si $\nu(M) > 1$, alors on peut également supposer grâce au lemme 2.37 que $\nu(m_{1,1}) = \nu(M)$. Cet élément jouera le rôle du pivot. S'il existe une ligne i telles que $m_{1,1}$ ne divise pas $m_{i,1}$, alors on a une division euclidienne :

$$m_{i,1} = q_i m_{1,1} + r_i$$

avec $r_i \neq 0$ et donc $\nu(r_i) < \nu(m_{1,1})$. En effectuant donc l'opération élémentaire :

$$L_i \longleftarrow L_i - q_i L_1$$

On obtient l'existence d'une matrice $P_1 \in SL_m(A)$ telle que :

$$P_1^{-1}M = \begin{pmatrix} m_{1,1} & \\ \vdots & \\ r_i & * \\ \vdots & \\ m_{m,1} & \end{pmatrix}$$

de sorte à ce que $\nu(P_1^{-1}M) < \nu(M)$ et donc en appliquant l'hypothèse de récurrence à $P_1^{-1}M$, on a le résultat escompté. Supposons donc que pour tout $i \in [\![2, m]\!]$, $m_{1,1}$ divise $m_{i,1}$. On a alors :

$$\forall i \in [\![2, m]\!],\ \exists q_i \in A, \quad m_{i,1} = q_i m_{1,1}.$$

En effectuant les opérations élémentaires :

$$\forall i \in [\![2, m]\!], \quad L_i \longleftarrow L_i - q_i L_1$$

on obtient l'existence d'une matrice $P_2 \in SL_m(A)$ telle que :

$$P_2^{-1}P_1^{-1}M = \begin{pmatrix} m_{1,1} & * \\ \mathbf{O}_{m-1,1} & * \end{pmatrix}.$$

On effectue les mêmes opérations sur les colonnes : s'il existe $j \in [\![2, n]\!]$ tel que $m_{1,1}$ ne divise pas $m_{1,j}$ alors en effectuant une division euclidienne :

$$m_{1,j} = q_j m_{1,1} + r_j$$

avec $\nu(r_j) < \nu(m_{1,1})$ et en effectuant l'opération élémentaire :

$$C_j \longleftarrow C_j - q_j C_1$$

on obtient, après multiplication à droite par une matrice $Q_1 \in SL_n(A)$, une matrice dont le coefficient en ligne 1 et colonne j est de stathme strictement plus petit que $\nu(M)$. On peut donc conclure en utilisant l'hypothèse de récurrence. On suppose alors que pour tout $j \in [\![2, n]\!]$, $m_{1,1}$ divise $m_{1,j}$. On a alors :

$$\forall j \in [\![2, n]\!],\ \exists q_j \in A, \quad m_{1,j} = q_j m_{1,1}.$$

En effectuant les opérations élémentaires :

$$\forall j \in [\![2, n]\!], \quad C_j \longleftarrow C_j - q_j C_1$$

on obtient l'existence d'une matrice $Q_2 \in SL_n(A)$ telle que :

$$P_2^{-1}P_1^{-1}MQ_1Q_2 = \begin{pmatrix} m_{1,1} & \mathbf{O}_{1,n-1} \\ \mathbf{O}_{m-1,1} & M' \end{pmatrix}.$$

ATTENTION ! Si vous appliquez l'hypothèse de récurrence à M', vous serez dans la **panade !** Car $m_{1,1}$ ne divisera pas forcément les facteurs invariants de M' !! On s'en sort ainsi : s'il existe $(i, j) \in [\![2, m]\!] \times [\![2, n]\!]$ tel que

$m_{1,1}$ ne divise pas $m_{i,j}$, alors, en effectuant l'opération élémentaire :

$$C_1 \longleftarrow C_1 + C_j$$

on se retrouve avec la matrice :

$$\begin{pmatrix} m_{1,1} & \mathbf{O}_{1,n-1} \\ m_{2,j} & \\ \vdots & M' \\ m_{n,j} & \end{pmatrix}$$

mais on a déjà traité ce cas-là ! En effet, il suffit de considérer une division euclidienne :

$$m_{i,j} = q_{i,j} m_{1,1} + r_{i,j}$$

et d'appliquer l'opération élémentaire :

$$L_i \longleftarrow L_i - q_{i,j} L_1$$

pour obtenir une matrice M'' dont le coefficient en ligne i et colonne 1 est $r_{i,j}$ et donc $\nu(M'') < \nu(M)$ et on conclut par récurrence ! Il ne reste plus qu'à supposer que $m_{1,1}$ divise $m_{i,j}$ pour tout couple $(i,j) \in [\![2,m]\!] \times [\![2,n]\!]$. Rappelons qu'on a :

$$P_2^{-1} P_1^{-1} M Q_1 Q_2 = \begin{pmatrix} m_{1,1} & \mathbf{O}_{1,n-1} \\ \mathbf{O}_{m-1,1} & M' \end{pmatrix}.$$

En appliquant l'hypothèse de récurrence à M' (ce qui est légitime car $t(M') < t(M)$) on a existence de matrices $(P', Q') \in SL_{m-1}(A) \times SL_{n-1}(A)$ et d'éléments $d_2 \mid \ldots \mid d_s \in A$ tels que :

$$P'^{-1} M' Q' = \begin{pmatrix} d_2 & 0 & \cdots & 0 & \\ 0 & d_3 & \cdots & 0 & \mathbf{O}_{s,n-1-s} \\ \vdots & \ddots & \ddots & \vdots & \\ 0 & \cdots & 0 & d_s & \\ & & & & \\ & \mathbf{O}_{m-1-s,s} & & & \mathbf{O}_{m-1-s,n-1-s} \\ & & & & \end{pmatrix}.$$

En notant alors :

$$P_3 = \begin{pmatrix} 1 & \mathbf{O}_{1,m-1} \\ \mathbf{O}_{m-1,1} & P' \end{pmatrix} \in SL_m(A),$$

$$Q_3 = \begin{pmatrix} 1 & \mathbf{O}_{1,n-1} \\ \mathbf{O}_{n-1,1} & Q' \end{pmatrix} \in SL_n(A),$$

et $d_1 := m_{1,1}$, on a :

$$P_3^{-1} P_2^{-1} P_1^{-1} M Q_1 Q_2 Q_3 = \begin{pmatrix} d_1 & 0 & \cdots & 0 & \\ 0 & d_2 & \cdots & 0 & \mathbf{O}_{s,n-s} \\ \vdots & \ddots & \ddots & \vdots & \\ 0 & \cdots & 0 & d_s & \\ & & & & \\ & \mathbf{O}_{m-s,s} & & & \mathbf{O}_{m-s,n-s} \\ & & & & \end{pmatrix}.$$

Or, les coefficients $d_2, \ldots, d_s$ sont calculés comme combinaisons A-linéaires des coefficients de la matrice extraite

M'. Étant donné que $d_1 = m_{1,1}$ divisait déjà tous les coefficients de cette matrice, on a donc, en particulier, $d_1 \mid d_2$. On a donc les relations de divisibilité :

$$d_1 \mid \ldots \mid d_s.$$

Cela termine donc l'hérédité, et donc la preuve de l'existence !

Unicité : La preuve de l'unicité repose sur les faits suivants :

Proposition 2.38 (Formules de Cauchy-Binet (voir [4]))**.** Soient K un corps et $(M, N) \in \mathscr{M}_{m,r}(K) \times \mathscr{M}_{r,n}(K)$. Pour $k \in [\![1, \min(m, n, r)]\!]$, on note $\mathbf{I}_{k,n}$ l'ensemble des applications strictement croissantes de $[\![1, k]\!]$ dans $[\![1, n]\!]$. Notons également, pour $(I, J) \in \mathbf{I}_{k,m} \times \mathbf{I}_{k,r}$ $m_{IJ}(M)$ le mineur de taille k associé au choix des lignes $I(1), \ldots, I(k)$ et au choix des colonnes $J(1), \ldots, J(k)$. On a alors :

$$\forall (I, J) \in \mathbf{I}_{k,m} \times \mathbf{I}_{k,n}, \quad m_{IJ}(MN) = \sum_{L \in \mathbf{I}_{k,r}} m_{IL}(M) m_{LJ}(N).$$

Cette identité peut se prolonger, au moyen de matrices universelles, aux matrices à coefficients dans un anneau commutatif unitaire quelconque ! On a alors le fait suivant :

Corollaire 2.39 (Invariance des idéaux de mineurs)**.** Soient A un anneau commutatif unitaire intègre et soient $M \in \mathscr{M}_{m,n}(A)$ et $P \in GL_m(A)$. On note $\mathscr{I}_k(M)$ l'idéal de A engendré par les mineurs de taille k de M. On a :

$$\forall k \in [\![1, \min(m, n)]\!], \quad \mathscr{I}_k(PM) = \mathscr{I}_k(M).$$

Démonstration. Par Cauchy-Binet, on a :

$$\mathscr{I}_k(PM) \subset \mathscr{I}_k(M) = \mathscr{I}_k(P^{-1}PM) \subset \mathscr{I}_k(PM),$$

ce qui conclut. Le résultat est aussi vrai lorsqu'on multiplie à droite par une matrice inversible. □

Ainsi, en notant D la forme normale de Smith obtenue dans notre partie existence, on obtient :

$$\forall k \in [\![1, s]\!], \quad \mathscr{I}_k(M) = \mathscr{I}_k(D) = \left(\prod_{i=1}^{k} d_i \right)$$

mais également :

$$\forall k \in [\![s+1, \min(m, n)]\!], \quad \mathscr{I}_k(M) = \mathscr{I}_k(D) = \{0\}.$$

Ainsi, les idéaux $(d_1), (d_1 d_2), \ldots (d_1 d_2 \ldots d_s)$ mais également l'entier s sont déterminés par M (s est le plus grand entier k tel que les mineurs de taille k engendrent un idéal non-nul). Ainsi, les idéaux $(d_1), (d_2), \ldots, (d_s)$ sont déterminés par M. En effet, en prenant, pour $k \in [\![1, s]\!]$, n'importe quel générateur Δ_k de $\mathscr{I}_k(M)$, on a que Δ_k divise Δ_{k+1} car $\mathscr{I}_{k+1}(M) \subset \mathscr{I}_k(M)$ et donc :

$$(d_k) = \left(\frac{\Delta_{k+1}}{\Delta_k} \right).$$

□