Une version vectorielle des bases standard

par Patrick Teller

L'Algèbre commutative a développé les bases de Grobner qui permettent d'étendre la notion de division polynomiale aux anneaux de polynômes en plusieurs indéterminées, de sorte que soit défini le reste d'un polynôme modulo un idéal; R. Thomas a introduit une interprétation géométrique du cas des binômes dans l'étude des applications à la programmation linéaire en entiers [1] mais il nous semble que la piste ouverte à cette occasion mérite d'être explorée plus avant.

L'idée de base consiste à représenter les binômes sans facteurs communs par des vecteurs, comme dans l'exemple suivant:

$\mathrm{x}_1^{\mathrm{n}_1}\ \mathrm{x}_2^{\mathrm{n}_2}\mathrm{x}_4^{n_4} - x_3^{n_3}x_5^{n_5}$ sera représenté par le vecteur $\begin{pmatrix} n_1 \\ n_2 \\ -n_3 \\ n_4 \\ -n_5 \end{pmatrix}$.

Comme les exposants des x_i sont nécessairement positifs il n'y a aucune ambiguité.

Reste à mettre en place les outils nécessaires à la réduction d'un vecteur par un autre (l'analogue de la division) et la réduction d'un vecteur par (modulo) une famille finie.

Dans tout ce qui suit « combinaison linéaire » s'entendra comme combinaison linéaire à coefficients dans $\mathbb{Z}$; de même on utilisera indistinctement les expressions p-uplet et vecteur pour des éléments de $\mathbb{Z}^p$.

1. Les outils (inspirés des bases de Grobner)

Nous désignerons par $\preccurlyeq$ la relation d'ordre (partiel) sur $\mathbb{N}^p$ définie comme suit $U = \begin{pmatrix} u_1 \\ ... \\ u_p \end{pmatrix} \preccurlyeq V = \begin{pmatrix} v_1 \\ ... \\ v_p \end{pmatrix} \Longleftrightarrow \forall i \in \{1, ..., p\}, u_i \leqslant v_i$; on dira alors que U précède V, lorsque 0 précède U on dira que U est positif.

On désignera par α l'ordre lexicographique qui, lui, est un ordre total; lorsque $\begin{pmatrix} 0 \\ . \\ 0 \end{pmatrix} \alpha \begin{pmatrix} v_1 \\ ... \\ v_p \end{pmatrix}$ on dira que $\begin{pmatrix} v_1 \\ ... \\ v_p \end{pmatrix}$ est lex-positif , de meme on dira lex-strictement-positif pour lex-positif et non nul.

Définition 1. *Soit un p-uplet* $U = \begin{pmatrix} u_1 \\ ... \\ .; \\ u_p \end{pmatrix}$ *on notera, pour chaque i,* $u_i^+ = max\{u_i, 0\}$ *et* $u_i^- = \max\{0, -u_i\}$, $U^+ = \begin{pmatrix} u_1^+ \\ ... \\ .; \\ u_p^+ \end{pmatrix}$ *et* $U^- = \begin{pmatrix} u_1^- \\ ... \\ .; \\ u_p^- \end{pmatrix}$, *d'où* $U = U^+ - U^-$; *lorsque* $U \neq 0$ *on notera* $s(U) = min\{i \in [\![1, ..., p]\!], u_i \neq 0\}$ *et* $U^* = U$ *lorsque* $u_{s(U)} > 0$ *et* $U^* = -U$ *lorsque* $u_{s(U)} < 0$; *si U est un p-uplet non nul* U^* *sera donc lex-strictement-positif.*

On appellera support de U l'ensemble $supp(U) = \{i, u_i \neq 0\}$ *et de même on pourra s'intéresser à* $supp(U^+)$ *et* $supp(U^-)$.

Par ailleurs on remarquera que si U est non nul alors U^* *est nécessairement lex-strictement positif.*

Définition 2. *Réduction d'un p uplet U par un p uplet V lex-strictement positif.*

Soit U et V deux p uplets tels que $\mathrm{V}^+ \preccurlyeq U^{*+}$, soit $\beta = \max\left\{k \in \mathbb{N},\ 0 \preccurlyeq U^{*+} - \mathrm{kV}^+ \right\}$, d'où on posera U=$\varepsilon\beta V + W$ (où $\varepsilon = 1$ si $\mathrm{U}^* = U, \varepsilon = -1$ si $\mathrm{U}^* = -U$). on remarquera (proposition 3) que V^+ ne précède plus ni W^+, ni W^-.

W sera appelé le reste (ou la réduction) de U modulo (par) V.

Dans le cas U=0V+W on dira que U est irréductible par V.

Proposition 3.

Soit la réduction de U par V U=$\gamma V + W$ alors V^+ ne précède ni W^+, ni W^-

Démonstration.

Il faut considérer deux cas

i) $U^{*+} - \beta V^+ = 0$, dans ce cas dans W les coordonnées appartenant au support de U^{*+} sont nulles, or $V^+ \preccurlyeq U^{*+}$ donc V^+ ne peut précéder ni W^+ ni W^-.

ii) $U^{*+} - \beta V^+$ n'est pas nul, mais comme $\beta = \max\left\{k \in \mathbb{N},\ 0 \preccurlyeq U^{*+} - \mathrm{kV}^+ \right\}$ V^+ ne peut précéder W^+; d'autre part, comme $U^{*+} - \beta V^+$ n'est pas nul il manque à W^+ et à W^- au moins une coordonnée pour être précèdé par V^+. □

Exemple 4.

U=$\begin{pmatrix} 4 \\ 3 \\ -3 \\ -2 \end{pmatrix}$, V=$\begin{pmatrix} 1 \\ 3 \\ -1 \\ -4 \end{pmatrix}$, U=1V+ W=$\begin{pmatrix} 3 \\ 0 \\ -2 \\ 2 \end{pmatrix}$, donc U $\overset{V}{\longmapsto} \begin{pmatrix} 3 \\ 0 \\ -2 \\ 2 \end{pmatrix}$

U=$\begin{pmatrix} -4 \\ -3 \\ 3 \\ 2 \end{pmatrix}$, V=$\begin{pmatrix} 1 \\ 3 \\ -1 \\ -4 \end{pmatrix}$, U=(-1)V+W=$\begin{pmatrix} -3 \\ 0 \\ 2 \\ -2 \end{pmatrix}$, donc U $\overset{V}{\longmapsto} \begin{pmatrix} -3 \\ 0 \\ 2 \\ -2 \end{pmatrix}$

U=$\begin{pmatrix} 0 \\ -3 \\ 3 \\ 2 \end{pmatrix}$, V=$\begin{pmatrix} 1 \\ 3 \\ -1 \\ -4 \end{pmatrix}$, U^*=$\begin{pmatrix} 0 \\ 3 \\ -3 \\ -2 \end{pmatrix}$ n'est pas précédé par V donc U $\overset{V}{\longmapsto} U$

Proposition 5. *Le processus de r*éduction *et les premières coordonnées*

Démonstration.

1) Soit U=$\begin{pmatrix} a \\ \cdots \\ \cdots \\ \cdots \end{pmatrix}$, où a>0; il ne peut être réduit que par un vecteur lex-strictement positif V=$\begin{pmatrix} b \\ \cdots \\ \cdots \end{pmatrix}$, où $0 \leqslant b \leqslant a$ et alors U $\overset{V}{\longmapsto} \begin{pmatrix} c \\ \cdots \\ \cdots \end{pmatrix}$, où $0 \leqslant c \leqslant a$.

2) Soit U=$\begin{pmatrix} 0 \\ a \\ \cdots \\ \cdots \end{pmatrix}$, où a>0; il ne peut être réduit que par un vecteur lex-strictement positif V=$\begin{pmatrix} 0 \\ b \\ \cdots \\ \cdots \end{pmatrix}$, où $0 \leqslant b \leqslant a$ et alors U $\overset{V}{\longmapsto} \begin{pmatrix} 0 \\ c \\ \cdots \end{pmatrix}$, où $0 \leqslant c \leqslant a$.

3) Soit U=$\begin{pmatrix} 0 \\ a \\ \cdots \\ \cdots \end{pmatrix}$, où a<0; alors U^*=$\begin{pmatrix} 0 \\ -a \\ \cdots \\ \cdots \end{pmatrix}$il ne peut être réduit que par un vecteur lex-strictement positif V=$\begin{pmatrix} 0 \\ b \\ \cdots \end{pmatrix}$, où $0 \leqslant b \leqslant -a$ et alors U $\overset{V}{\longmapsto} \begin{pmatrix} 0 \\ c \\ \cdots \end{pmatrix}$. □

Proposition 6.

Le reste d'un vecteur positif modulo un vecteur lex-strictement positif est positif

Démonstration.

Soit $U=\begin{pmatrix} u_1 \\ \cdots \\ \cdots \\ u_p \end{pmatrix} \succcurlyeq 0$ *et* $V=\begin{pmatrix} v_1 \\ \cdots \\ \cdots \\ v_p \end{pmatrix}$, lex − strictement positif*;* si U n'est pas irréductible alors $\forall i$, $v_i > 0$,$u_i \geqslant v_i$ et le reste de U par V sera U-max$\{u_i/v_i, v_i>0\}$V, donc à coordonnées positives. □

Définition 7. *Réduction d'un p uplet U par un ensemble de p uplets lex-strictement positifs S*

Soit le n+m–uplet U et un ensemble de p -uplets $S=\{V_i, i \in I\}$ *on dira que U est réduit à W modulo S lorsqu'il existe un ensemble d'indices* $(i_1,..,i_q)$ *à valeurs dans I tel que* $U \overset{V_{i_1}}{\longmapsto} \overset{V_{i_2}}{\longmapsto} ..... \overset{V_{i_q}}{\longmapsto} W$, *où W est irréductible par les différents* V_i*; W pourra être appelé « un reste » de U modulo S.*

Avertissement 8.

Dans le cas general le reste modulo un ensemble de p uplets n'est pas unique, comme le prouve l'exemple suivant

$U=\begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix}$, $V_1=\begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$, $V_2=\begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$

$U \overset{V_1}{\longmapsto} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \overset{V_2}{\longmapsto} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ tandis que $U \overset{V_2}{\longmapsto} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \overset{V_1}{\longmapsto} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$

Rappelons le

Lemme 9. *Gordon-Dickson*

Toute partie non vide de N^p *possède un nombre fini d'éléments minimaux (pour la relation* $\preccurlyeq$*)*

(généralise le classique minimum d'une partie non vide de $\mathcal{N}$; pour les algébristes: ce résultat nous épargne les recours Noetheriens)

Théorème 10. *(d'après le théorème de Buchberger)*

On suppose donné un ensemble fini de p-uplets lex-strictement positifs $S=\{V_i, i \in I\}$.

On pose $S:=\{V_i, i \in I\}$, $G:=\{\{V_i, V_j\}\}, V_i \neq V_j\}$

tant que $G \neq \varnothing$

on prend $\{V_i, V_j\} \in G$, *on pose* $G:=G-\{V_i, V_j\}$,*W:=une réduction de* $(V_i - V_j)$ *modulo S*

si $W \neq 0$ *on pose* $G:=G \cup \{\{V, W^*\}, V \in G\}, S := S \cup \{W^*\}$

L'algorithme de Buchberger s'arrête en un temps fini.

Remarquons au passage que, suite à l'hypothèse sur les V_i *et la définition des* W^**, les éléments de S sont tous lex-positifs.*

Démonstration.

Désignons par (S_k) *la suite des ensembles S créés par l'algorithme et, pour chaque k, par* S_k^+ *l'ensemble* $\{V^+, V^-\}_{V \in S_k}$*; tant qu'on obtient un p uplet W* avec W*+ et/ou W*− non nul* $S_k^+ \subsetneq S_{k+1}^+$*;* $S_\infty^+ = \cup_{k\in\mathbb{N}} S_k^+$ *est une partie non vide de* $\mathbb{N}^n$ *et possède donc un ensemble fini d'éléments minimaux pour* $\preccurlyeq$*, par suite il existe un indice k tel que* S_k^+ *contient tous ces éléments minimaux; or pour tout élément W* créé par l'algorithme ni W*+ ni W*− n'est précédé par aucun élément préexistant (proposition 3), par suite à partir d'un certain moment les W*+ et les W*− sont nuls, par suite W*=0 et l'algorithme de Buchberger s'arrête donc* $S_\infty = S_k$ *et sera appelé (une) base standard associée à S.*

On retiendra que les éléments de la base standard sont lex-strictement positifs, que S_∞ *contient S et que les ensembles S et* S_∞ *engendrent le même sous-groupe additif.*

□

Exemple 11.

On considère les trois uplets $\begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 0 \\ 0 \\ -1 \end{pmatrix}$, l'application de l'algorithme produit les six uplets: $\begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 7 \\ -1 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -5 \\ 1 \\ 1 \end{pmatrix}$.

Lemme 12.

Soit la base standard $S_\infty = \{V_i, i \in I\}$*, où* $\forall i, V_i = \begin{pmatrix} v_i \\ \dots \\ \dots \end{pmatrix}$.

1) Soit $V = \sum_{i\in I} m_i V_i$ *(où les* m_i *sont dans* $\mathbb{Z}$*) et* $\sum_{i\in I} m_i v_i = 0$ *, alors il existe un sous-ensemble* $\mathcal{G} = \{W_i, i \in I'\} \subset S_\infty$ *tel que*

$V = \sum_{i\in I'} n_i W_i$ *et* $\forall i \in I', W_i = \begin{pmatrix} 0 \\ \dots \\ \dots \end{pmatrix}$

2) Soit $V = \sum_{i\in I} m_i V_i$ *et* $\sum_{i\in I} m_i v_i \neq 0$*, alors il existe un élément W de* S_∞ *tel que le reste de V modulo W est une combinaison linéaire d'éléments de* S_∞ *dont la première coordonnée est nulle.*

Nous allons utiliser la proposition 7. qui signifie que pour tout couple (V_i,V_j) d'éléments de la base standard $(V_i - V_j)^* = \begin{pmatrix} a \geqslant 0 \\ .... \\ ... \end{pmatrix}$ est une combinaison linéaire d'éléments de la base standard, et pour chacun de ceux-ci la première coordonnée appartient à [0,a].

1) Soit V=$\sum_{i\in I} m_i V_i$ tel que $\sum_{i\in I} m_i v_i = 0$ et K=$\{i, s(V_i) = 1\}$ n'est pas vide.

Pour chacun des i∈K nous savons v_i est strictement positif, et les coefficients m_i sont quant à eux des entiers relatifs et la somme $\sum_{i\in K} m_i v_i = 0$.

Montrons qu'il est possible pas à pas de diminuer le cardinal de K.

Soit $K^- = \{k \in K, m_k < 0\}$ et $K^+ = \{k \in K, m_k > 0\}$, comme $\sum_{i\in K} m_i v_i = 0$ ni K^- ni K^+ ne peuvent être vides.

On pourra tout au long de la procédure suivante, quitte à introduire des différences $(V_i - V_j)^*$ dont la première coordonnée serait alors nulle, toujours supposer que $\forall (i,j) \in K^- \times K^+, v_i \neq v_j$

Soit $v_t = \max\{v_k, k \in K\}$

i) si c'est $v_j \in K^+$ on considère $v_s = \max\{v_i, i \in K^-\}$, alors $v_j - v_s > 0$

$m_s V_s + m_j V_j = m_j(V_j - V_s) + (m_s - m_j)V_s$; la proposition 5 permet de conclure que le $\max\{v_k, k \in K\}$ diminue strictement

ii) si c'est $v_i \in K^-$ on considère $v_t = \max\{v_j, j \in K^+ 0\}$, alors $v_i - v_t > 0$

$m_i V_i + m_t V_t = m_i(V_i - V_t) + (m_i + m_t)V_t$ d'où, de même le $\max\{v_k, k \in K\}$ diminue strictement

comme les v_k sont des entiers naturels ce processus sera fini.

Comme nous avons remarqué que, si K n'est pas vide, ni K^- ni K^+ne peuvent l'être, nous aboutirons en fin de compte à $K = \varnothing$, ce qui démontre le résultat annoncé: nous avons obtenu une écriture de V comme combinaison linéaire d'éléments de la base standard tels que $s > 0$.

2) Soit $V = \sum_{i \in I} m_i V_i$ et $\sum_{i \in I} m_i v_i = v \neq 0$, c'est à dire $\sum_{i \in I} m_i v_i = v > 0$, puisque V est lex-positif.

Soit $v_t = \max\{v_k, k \in I\}$

En appliquant la démarche du 1) on peut supposer que les coefficients m_i sont tous strictement positifs et, quitte à introduire des différences $(V_i - V_j)^*$ dont la première coordonnée serait alors nulle, toujours supposer que $\forall (i,j) \in I^2, i \neq j \Longrightarrow v_i \neq v_j$.

Si card(I)=1 V est la somme de $m_1 V_1$ et T, somme d'éléments de S_∞ dont la première coordonnée est nulle, donc $V = m_1 V_1 + T$ est une réduction.

Si card(I)>1 considérons les deux indices i,j qui correspondent aux deux plus grandes valeurs de $\{v_k\}$:

supposons $v_j < v_i$ $m_i V_i + m_j V_j = m_i(V_i - V_j) + (m_i + m_j)V_j$, où $V_i - V_j$ est une combinaison linéaire de termes lex-positifs de la base standard dont le premier indice est strictement inférieur à v_i

En itérant on obtient en fin de compte $V = mW + T$, où W est un élément de S_∞ et T une somme d'éléments de S_∞ dont la première coordonnée est nulle.

Théorème 13.

Soit un ensemble fini de n+m uplets $S = \{V_i, i \in I\}$ et une base standard $S_\infty = \{W_j, j \in J\}$ associée à S; un n+m-uplet V appartient au sous-groupe engendré par S si et seulement il est réduit modulo S_∞.

Démonstration.

Soit $V = \sum a_j W_j$

1) Dans un premier temps nous allons montrer que, quel que soit $V = \sum_{j \in J} m_j W_j$, où les m_j sont des entiers relatifs non nuls, il existe une écriture de ce type de V, où s(V) est égal au minimum des $s(W_j)$.

Il est clair que, quelle que soit la décomposition considérée, $s(V) \geqslant \min(\{s(W_j), j \in J\})$, il suffit donc de montrer que, V étant choisi, il existe une décomposition telle que $s(V) \leqslant \min(\{s(W_j), j \in J\})$.

Supposons, pour alléger l'écriture, que $s(V) \geqslant 1$ et qu'il existe une décomposition $V = \sum_{j \in J} m_j W_j$ où $K = \{j, s(W_j) = 0\}$ n'est pas vide.

Le lemme 12 montre qu'on peut trouver une écriture de $V = \sum_{j \in J} m_j W_j$, comme une combinaison lineaire d'éléments de la base standard tels que $s(W_j) = 0$; en répétant on obtient une décomposition, $V = \sum_{s(W_j) = s(V)} m_j W_j + \sum_{s(W_j) > s(V)} m_j W_j$.

La seconde partie du lemme montre que ceci est le début du processus de réduction de V.

Par suite V est réduit modulo la base standard. □

Théorème 14. *Unicité du reste d'un p -uplet modulo S_∞*

Soit V un p-uplet, le reste de V dans une réduction modulo S_∞ est unique (indépendant de la procédure utilisée).

Démonstration.

Soit V un n+m-uplet, supposons que $\begin{cases} V = \sum_{j\in J} m_j W_j + R_1 \\ V = \sum_{j\in J} \mu_j W_j + R_2 \end{cases}$ *deux réductions de V modulo S_∞, alors (R_1-R_2)* appartient au sous-groupe engendré par les W_j et donc son reste est nul; par ailleurs (R_1-R_2)* n'est précédé par aucun des W_j donc il est égal à son reste.*

Donc le reste de V modulo S_∞ est indépendant de la réduction opérée. □

Comparaison avec les bases standard « classiques »

Dans le cas des binômes il semble que la forme proposée est plus légère:

Au lieu d'un binôme $a^2x^3y - b^0z^3t^2$ nous écrirons $\begin{pmatrix} 2\\0\\3\\1\\-3\\-2 \end{pmatrix}$ et au lieu de ay^2z-$b^2x^2t^3$ nous écrirons $\begin{pmatrix} 1\\-2\\-2\\2\\1\\-3 \end{pmatrix}$ (avec l'ordre a>b>x>y>z>t).

Le calcul du S-polynome associé passe par le ppcm des monômes a^2x^3y et ay^2z qui est $a^2x^3y^2z$, puis yz($a^2x^3y - b^0z^3t^2$)-ax^3($ay^2z - b^2x^2t^3$)=-$a^0b^0yz^4t^2$+$ab^2x^5t^3$=t^2(-$a^0b^0yz^4t^0$+$ab^2x^5z^0t$), où l'on simplifie par le facteur commun en invoquant les propriétés de l'idéal d'un treillis, d'où le vecteur $\begin{pmatrix} 1\\2\\5\\-1\\-4\\1 \end{pmatrix}$.

Tandis qu'en soustrayant simplement les vecteurs on obtient directement $\begin{pmatrix} 2\\0\\3\\1\\-3\\-2 \end{pmatrix} - \begin{pmatrix} 1\\-2\\-2\\2\\1\\-3 \end{pmatrix} = \begin{pmatrix} 1\\2\\5\\-1\\-4\\1 \end{pmatrix}$.

La présentation proposée évite l'usage de polynômes en plusieurs indéterminées et des connaissances algébriques associées, permettant un emploi par un public d'étudiants beaucoup plus large.

Bibliographie :[1] R.R. Thomas (1995), A geometric Buchberger Algorithm for integer programming, Math.Operations Research 20, pp.864-884

Paris, Mars 2017