Tout polytope rationnel a un système d'équations dans $\mathbb{N}$

par PATRICK TELLER

Complété le 26 décembre 2019

Il est évident que l'ensemble des solutions du système diophantien $\begin{cases} \sum_{j=1}^{n} a_j x_j = b \\ (x_1, ..., x_n) \geqslant 0 \end{cases}$ est fini et non vide si et seulement si il est équivalent à un système de la forme $\begin{cases} \sum_{j=1}^{n} a_j x_j = b \\ (x_1, ..., x_n) \geqslant 0 \end{cases}$, où le n+1 uplet $(a_1, ..., a_n, b)$ appartient à $\mathbb{N}^{*n} \times \mathbb{N}$; nous allons établir un résultat analogue pour le cas d'un système $\begin{cases} \mathrm{AX} = b \\ X \geqslant 0 \end{cases}$.

Pour celà nous allons appliquer l'algorithme du criss-cross [3] qui est une des méthodes primales-duales du simplexe; le premier paragraphe a été rédigé pour permettre la lecture sans nécessité de familiarité avec le monde de la programmation linéaire et des techniques du simplexe.

La preuve de la faisabilité (c'est à dire l'assurance d'aboutir en un « temps fini ») de l'algorithme utilisé sera donnée; elle s'inspire de [4] mais elle est plus élémentaire, s'appuyant sur le caractère non vide et borné de P.

1 Définitions et Préparation

Définition 1. *La matrice étendue d'un système d'équations diophantiennes* $\begin{cases} \mathrm{AX} = b \\ X \geqslant 0 \end{cases}$

Soient (A,b)$\in \mathcal{M}_{\mathrm{mn}}(\mathbb{Z}) \times \mathcal{M}_{m1}(\mathbb{Z})$, on appelle matrice étendue du système $\begin{cases} \mathrm{AX} = b \\ X \geqslant 0 \end{cases}$ la matrice (A,b)$\in \mathcal{M}_{m,n+1}(\mathbb{Z})$; si P est l'ensemble des solutions du système $\begin{cases} \mathrm{AX} = b \\ X \geqslant 0 \end{cases}$ on parlera de matrice étendue du polytope P.

Lemme 2. *(préparatoire)*

Soient (A,b)$\in \mathcal{M}_{m,n}(\mathbb{Z}) \times \mathcal{M}_{m,1}(\mathbb{Z})$ *et le système* $\begin{cases} \mathrm{AX} = b \\ X \geqslant 0 \end{cases}$ *,où rang(A)=m.*

Le système $\begin{cases} \mathrm{AX} = b \\ X \geqslant 0 \end{cases}$ *est équivalent à un système de la forme* $\begin{cases} A''X = b'' \\ X \geqslant 0 \end{cases}$ *, à coefficients dans* $\mathbb{Q}$*, où la matrice A" est, à l'ordre des colonnes près, de la forme* $\begin{pmatrix} A_1 & I_{m-1} \\ L_1 & 0 \end{pmatrix}$*,avec* $A_1 \in \mathcal{M}_{m-1,n-m+1}(\mathbb{Q})$ *et* $L_1 \in \mathcal{M}_{1,n-m+1}(\mathbb{Q})$*.*

Démonstration. □

On pose $A = \begin{pmatrix} A' \\ L \end{pmatrix}$*, où* $A' \in \mathcal{M}_{m-1n}(\mathbb{Z})$ *et* $L \in \mathcal{M}_{1n}(\mathbb{Z})$ *et* $b = \begin{pmatrix} b' \\ b_m \end{pmatrix}$

A étant de rang m A' est de rang m-1, donc il existe un sous-ensemble $B = \{a_1 < a_2 < .. < a_{m-1}\}$ *de {1,..,n} tel que les colonnes de A' qui portent ces indices forment une base de l'espace des colonnes de A'; on désigne par N le complémentaire de B dans {1,..,n}.*

Soient A'_B et A'_N les matrices extraites de A' associées à la partition de l'ensemble des colonnes en B et N.

$AX{=}b \Longleftrightarrow \begin{cases} A'X = b' \\ \mathrm{LX} = b_m \end{cases} \Longleftrightarrow \begin{cases} {A'_B}^{-1}A'X = {A'_B}^{-1}b' \\ \mathrm{LX} = b_m \end{cases}$ *; dans la matrice* ${A'_B}^{-1}A'$ *les colonnes dont les indices appartiennent à B sont de la forme* $\begin{pmatrix} 0 \\ .... \\ 0 \\ 1 \\ 0 \\ .... \\ 0 \end{pmatrix}$*, donc en retranchant à la dernière équation de ce système une bonne combinaison linéaire des autres le système devient* $\begin{cases} {A'_B}^{-1}A'X = {A'_B}^{-1}b' \\ \mathrm{CX} = c_m \end{cases}$ *,où les coordonnées de C dont les indices appartiennent à B sont nulles.*

Définition 3. *Tableau standard*

On appellera tableau standard de taille $m \times (n+1)$ toute matrice $M{=}(m_{\mathrm{ij}}){\in}\mathcal{M}_{\mathrm{mn+1}}(\mathbb{Q})$ de rang m qui vérifie les propriétés suivantes:

i) Si on désigne par M' la matrice M privée de sa dernière ligne et de sa dernière colonne, il existe un sous-ensemble $B{=}\{a_1{<}a_2{<}..{<}a_{m-1}\}$ de {1,..,n} tel que les colonnes de M' qui portent ces indices sont, à l'ordre près, les colonnes de la matrice unité; on désignera par j(i) l'indice de la colonne de M' dont la i-ème composante est un 1 et les m-2 autres sont nulles.

ii) $\forall j \in B$, $m_{\mathrm{mj}}{=}0$

Proposition 4.

La procédure décrite dans le lemme ci-dessus transforme le système d'équations d'un polytope en un système équivalent dont la matrice étendue est un tableau standard.

Démonstration.

Tout est dans le lemme. □

Définition 5. *Pivotage*

Soit une matrice $M{=}(m_{\mathrm{ij}}){\in}\mathcal{M}_{\mathrm{pq}}(\mathbb{Q})$, dont on notera les lignes M_i (i=1,..,p) et le couple $(i_0,j_0){\in}\{1,...,p\}\times\{1,...,q\}$ tel que $m_{i_0j_0} \neq 0$, on appellera pivotage par rapport à la case (i_0,j_0) la suite d'opérations suivantes:

$M_{i_0}{:}M_{i_0}/m_{i_0j_0}$

$\forall i \neq i_0, M_i{:}\, M_i - m_{ij_0}M_{i_0}$

Proposition 6.

Soit un polytope P et une matrice étendue $M{=}(m_{\mathrm{ij}}){\in}\mathcal{M}_{\mathrm{mn}}(\mathbb{Q})$ de P, si on considère un couple $(i_0,j_0){\in}\{1,...,m-1\}\times\{1,...,n\}$ tel que $i_0 \in N$ et $m_{i_0j_0} \neq 0$, l'image M^* de M, après pivotage par rapport à la case (i_0,j_0), est un tableau standard.

Démonstration.

On précisera d'abord, pour mémoire, que M^* est la matrice étendue d'un système d'équations de P.

Pour établir que M^* est un tableau standard il suffit de remarquer que

i) le pivotage préserve les m-1 premières coordonnées des colonnes d'indice j∈B sauf j(i_0)

*ii) le pivotage transforme la colonne d'indice j_0 en une colonne dont les m-1 premières coordonnées sont nulles sauf la i_0-*ième *qui devient un 1 (j_0« remplace » j(i_0))*

iii) les 0 de la dernière ligne qui se trouvent dans les colonnes d'indice j∈B sauf j(i_0) sont préservés parce que 0-0=0

iv) un 0 apparaît dans la dernière ligne-colonne d'indice j_0. □

2 Le Théorème

Nous allons établir le Théorème suivant:

Théorème 7.

Soient (A,b)∈$\mathcal{M}_{\mathrm{mn}}(\mathbb{Z})\times\mathcal{M}_{m1}(\mathbb{Z})$, où rang(A)=m, l'ensemble des solutions du système $\begin{cases} \mathrm{AX}=b \\ X\geqslant 0 \end{cases}$ *est borné si et seulement il existe une matrice T∈$\mathcal{GL}m(\mathbb{Z})$ telle que T(A,b)∈$\mathcal{M}_{m,n+1}(\mathbb{N})$.*

Autrement dit

Si on considère un polytope (=polyèdre non vide et borné) défini par le système d'équations $\begin{cases} \mathrm{AX}=b \\ X\geqslant 0 \end{cases}$, où (A,b)∈$\mathcal{M}_{\mathrm{mn}}(\mathbb{Z})\times\mathcal{M}_{m1}(\mathbb{Z})$, et rang(A)=m, il existe un système équivalent de la forme $\begin{cases} \mathrm{AX}=b \\ X\geqslant 0 \end{cases}$ où (A,b)∈$\mathcal{M}_{\mathrm{mn}}(\mathbb{N})\times\mathcal{M}_{m1}(\mathbb{N})$.

Nous allons utiliser une des versions de l'algorithme criss-cross [3].

Notation 8.

Soit la matrice étendue $M=\begin{pmatrix} M' & b' \\ C & c_m \end{pmatrix}$.

Pour chaque i∈B on désignera par b'$_i$ la coordonnée correspondante du vecteur b′.

De même on désigne par (s_1,..,s_n) le vecteur-ligne C.

Description de l'algorithme:

Etape 1:

Soit I={j(i),b'$_i$ <0} et J={j⩽n,s_j<0}

Si $I\cup J=\varnothing$, fin.

sinon, soit k=min(I∪J), si k∈I, aller en 2., si k∈J, aller en 3.

Etape 2:

Soit S={j, a_{kj}<0}

Si S=∅, le système est inconsistant,fin.

Sinon soit $j_0=\min(S)$, B:B∪$j_0\backslash k$. (et donc pivoter)

Etape 3:

Soit T={i, a_{ik}>0}

Si T=$\varnothing$, l'ensemble défini par AX=b n'est pas borné,fin.

Sinon soit i_0=min (T), B:B$\cup i_0 \backslash k$. (et donc pivoter)

Retourner en 1.

THEOREME 2:

Soient (A,b)$\in \mathcal{M}_{mn}(\mathbb{Z}) \times \mathcal{M}_{m1}(\mathbb{Z})$, où rang(A)=m, l'ensemble des solutions du système $\begin{cases} AX = b \\ X \geqslant 0 \end{cases}$ est borné si et seulement il existe une matrice T$\in \mathcal{GL}m(\mathbb{Q})$ telle que T(A,b)$\in \mathcal{M}_{m,n+1}(\mathbb{N})$

Démonstration.

L'algorithme décrit au-dessus s'arrête en un temps fini (voir paragraphe 5).

Soit A"X=b" le système obtenu où A"=(a"$_{ij}$)

Alors les s_j $(j \leqslant n)$sont positifs au sens large de même que les b$_i$",i<m.

Quant à b"$_m$ il est positif car il existe au moins une solution X à valeurs dans $\mathbb{Q}^{+n}$ et b"$_m$=$\sum s_j x_j$.

On désigne par $L_1, ..., L_m$ les équations du système équivalent obtenu au-dessus.

Si les s_j sont tous strictement positifs il suffit d'ajouter à chaque équation L_i un bon multiple de L_m.

Sinon

la clé est dans le fait que le caractère borné de P entraîne que quel que soit j il existe i tel que a"$_{ij}$>0.

Etape 1:

Soit Z={j,s_j=0$\wedge \exists i, a''_{ij} < 0$} et j$_1$=min(Z),I$_1$=$\{i, a''_{ij_1} < 0\}$ et i$_1$ tel que a"$_{i_1 j_1}$>0, alors tant que Z$\neq\varnothing$,

Etape 2:

A chaque équation L_i,i$\in$I$_1$ on ajoute un bon multiple de L_{i_1} et retour à 1.

Une fois acquis le fait que les coefficients sont positifs, il suffit de multiplier A" et b" par le plus petit dénominateur commun pour que tous les coefficients soient dans $\mathbb{N}$.

□

3 La preuve de l'algorithme

D'après Fukuda et Matsuy [4].

Comme nous avons supposé que P était à la fois borné et non vide la démonstration originale [4] pourra être largement allégée.

Nous supposerons, pour cette démonstration, que dans l'algorithme, après pivotage, l'ensemble B est réécrit de telle sorte que, comme au début il soit dans l'ordre naturel des entiers.

On considérera le domaine D défini par $\begin{cases} A'X = b' \\ X \geqslant 0 \end{cases}$ et la fonction f:X->-LX+c$_m$.

Dans les tableaux ci-dessous qui obéissent à l'algorithme Criss-Cross il faut interpréter - comme <0, $\boxminus$ comme $\leqslant 0$, + comme >0 et $\boxplus$ comme $\geqslant 0$.

Dans cette partie on supposera que le lecteur est familiarisé avec les notions élémentaires de la méthode du simplexe.

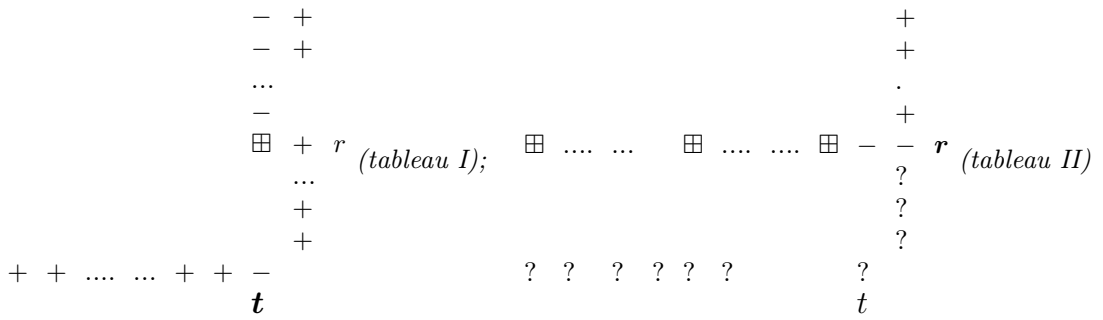
Proposition 9. *L'algorithme Criss-Cross aboutit en un temps fini*

Démonstration.

On va supposer qu'il existe un cycle et on désigne par F l'ensemble des indices qui entrent et qui sortent de B et soit t=max(F).

Il existe donc un indice r qui quitte B en faveur de t et un indice s qui remplace t lorsqu'il quitte B.

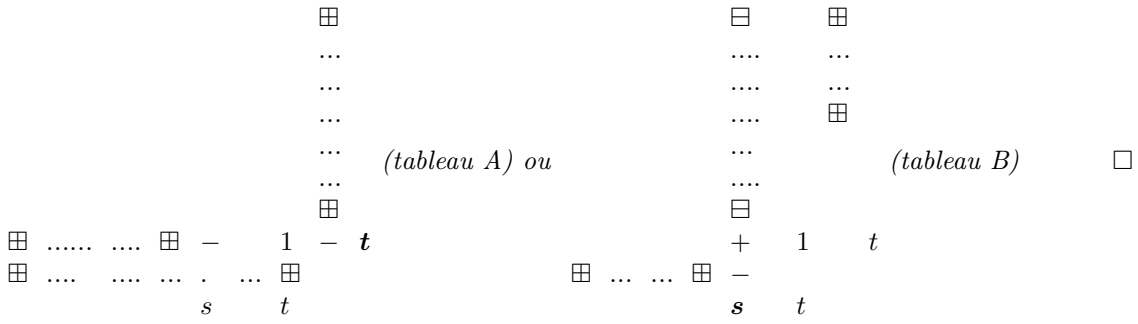
Voici les deux tableaux standards possibles juste avant de l'échange r<->t



Le tableau I nous apprend que la dérivée partielle de f par rapport à x_t est strictement positive.

La ligne r du tableau II se lit $\sum, a_{\mathrm{rj}}x_j + a_{\mathrm{rt}}x_t = b'_r$ et au point de base considéré $x_t=0$ puisque x_t est alors une variable hors base; donc $\sum, a_{\mathrm{rj}}x_j + 0 = b'_r$, or les coefficients a_{rj} sont positifs et $b'_r<0$ donc il n'y a pas de point de D où $x_t=0$.

Voici les deux tableaux standards possibles juste avant l'échange t<->s



Le tableau A se place en un point de base où x_t (de base) est strictement négatif et z maximal.

Le tableau B montre que x_s et xt varient de manière opposée et que la dérivée de f par rapport à xs est positive; on peut aussi dire que, si $x_t=0$, x_s n'est pas majoré.

I et A sont incompatibles.

De même I et B.

II établit qu'il n'y a pas de point où $x_t<0$ et les autres $x_j \geqslant 0$; A se place en un tel point, d'où l'incompatibilité.

Reste II et B: Si on considère le système privé de l'équation d'indice t et de l'inconnue x_t II signifie qu'il est inconsistant et B qu'il admet un ensemble de solutions non borné.

Par suite l'hypothèse sur x_t est absurde.

Exemple 10.

Soit le polygone P défini par le système $\begin{cases} 5x1+3x2\leqslant 45\\ 2x1-3x2\leqslant -9\\ -9x1+2x2\leqslant 4\\ 3x1+5x2\leqslant 50\\ (x1,x2)\geqslant 0 \end{cases}$, que nous écrirons sous la forme

```
Maxima 5.41.0 http://maxima.sourceforge.net
using Lisp GNU Common Lisp (GCL) GCL 2.6.12
Distributed under the GNU Public License. See the file COPYING.
Dedicated to the memory of William Schelter.
The function bug_report() provides bug reporting information.
```

```
(%i1) M:matrix([5,3,1,0,0,0,45],[2,-3,0,1,0,0,-9],[-9,2,0,0,1,0,4],[3,5,0,0,0,1,
      50]);
```

$$(\%o1)\quad \begin{pmatrix} 5 & 3 & 1 & 0 & 0 & 0 & 45\\ 2 & -3 & 0 & 1 & 0 & 0 & -9\\ -9 & 2 & 0 & 0 & 1 & 0 & 4\\ 3 & 5 & 0 & 0 & 0 & 1 & 50 \end{pmatrix}$$

```
(%i2) for j:1 thru 7 do M[2,j]:M[2,j]/2;M;
```

(%o2) **done**

$$(\%o3)\quad \begin{pmatrix} 5 & 3 & 1 & 0 & 0 & 0 & 45\\ 1 & -\frac{3}{2} & 0 & \frac{1}{2} & 0 & 0 & -\frac{9}{2}\\ -9 & 2 & 0 & 0 & 1 & 0 & 4\\ 3 & 5 & 0 & 0 & 0 & 1 & 50 \end{pmatrix}$$

```
(%i4) for i:3 thru 4 do (coeff:M[i,1],for j:1 thru 7 do M[i,j]:M[i,j]-coeff*M[2,
      j]);for j:1 thru 7 do M[1,j]:M[1,j]-5*M[2,j];M;
```

(%o4) **done**

(%o5) **done**

$$(\%o6)\quad \begin{pmatrix} 0 & \frac{21}{2} & 1 & -\frac{5}{2} & 0 & 0 & \frac{135}{2}\\ 1 & -\frac{3}{2} & 0 & \frac{1}{2} & 0 & 0 & -\frac{9}{2}\\ 0 & -\frac{23}{2} & 0 & \frac{9}{2} & 1 & 0 & -\frac{73}{2}\\ 0 & \frac{19}{2} & 0 & -\frac{3}{2} & 0 & 1 & \frac{127}{2} \end{pmatrix}$$

```
(%i7) for j:1 thru 7 do M[2,j]:-2*M[2,j]/3;M;
```

(%o7) **done**

$$(\%o8)\quad \begin{pmatrix} 0 & \frac{21}{2} & 1 & -\frac{5}{2} & 0 & 0 & \frac{135}{2}\\ -\frac{2}{3} & 1 & 0 & -\frac{1}{3} & 0 & 0 & 3\\ 0 & -\frac{23}{2} & 0 & \frac{9}{2} & 1 & 0 & -\frac{73}{2}\\ 0 & \frac{19}{2} & 0 & -\frac{3}{2} & 0 & 1 & \frac{127}{2} \end{pmatrix}$$

```
(%i9) for i:3 thru 4 do (coeff:M[i,2],for j:1 thru 7 do M[i,j]:M[i,j]-coeff*M[2,
      j]);for j:1 thru 7 do M[1,j]:M[1,j]-21*M[2,j]/2;M;
```

(%o9) **done**

(%o10) **done**

(%o11)
$$\begin{pmatrix} 7 & 0 & 1 & 1 & 0 & 0 & 36 \\ -\frac{2}{3} & 1 & 0 & -\frac{1}{3} & 0 & 0 & 3 \\ -\frac{23}{3} & 0 & 0 & \frac{2}{3} & 1 & 0 & -2 \\ \frac{19}{3} & 0 & 0 & \frac{5}{3} & 0 & 1 & 35 \end{pmatrix}$$

```
(%i12) for j:1 thru 7 do M[3,j]:-3*M[3,j]/23;M;
```

(%o12) **done**

(%o13)
$$\begin{pmatrix} 7 & 0 & 1 & 1 & 0 & 0 & 36 \\ -\frac{2}{3} & 1 & 0 & -\frac{1}{3} & 0 & 0 & 3 \\ 1 & 0 & 0 & -\frac{2}{23} & -\frac{3}{23} & 0 & \frac{6}{23} \\ \frac{19}{3} & 0 & 0 & \frac{5}{3} & 0 & 1 & 35 \end{pmatrix}$$

```
(%i14) for i:1 thru 2 do (coeff:M[i,1],for j:1 thru 7 do M[i,j]:M[i,j]-coeff*M[3,
       j]);for j:1 thru 7 do M[4,j]:M[4,j]-19*M[3,j]/3;M;
```

(%o14) **done**

(%o15) **done**

(%o16)
$$\begin{pmatrix} 0 & 0 & 1 & \frac{37}{23} & \frac{21}{23} & 0 & \frac{786}{23} \\ 0 & 1 & 0 & -\frac{9}{23} & -\frac{2}{23} & 0 & \frac{73}{23} \\ 1 & 0 & 0 & -\frac{2}{23} & -\frac{3}{23} & 0 & \frac{6}{23} \\ 0 & 0 & 0 & \frac{51}{23} & \frac{19}{23} & 1 & \frac{767}{23} \end{pmatrix}$$

```
(%i17) M[2]:M[2]+9*M[4]/51;M[3]:M[3]+3*M[4]/19;M;
```

(%o17) $\left[0, 1, 0, 0, \frac{1}{17}, \frac{3}{17}, \frac{154}{17}\right]$

(%o18) $\left[1, 0, 0, \frac{5}{19}, 0, \frac{3}{19}, \frac{105}{19}\right]$

(%o19)
$$\begin{pmatrix} 0 & 0 & 1 & \frac{37}{23} & \frac{21}{23} & 0 & \frac{786}{23} \\ 0 & 1 & 0 & 0 & \frac{1}{17} & \frac{3}{17} & \frac{154}{17} \\ 1 & 0 & 0 & \frac{5}{19} & 0 & \frac{3}{19} & \frac{105}{19} \\ 0 & 0 & 0 & \frac{51}{23} & \frac{19}{23} & 1 & \frac{767}{23} \end{pmatrix}$$

```
(%i20) M[1]:23*M[1];M[2]:17*M[2];M[3]:19*M[3];M[4]:23*M[4];M;
```

(%o20) $[0, 0, 23, 37, 21, 0, 786]$
(%o21) $[0, 17, 0, 0, 1, 3, 154]$
(%o22) $[19, 0, 0, 5, 0, 3, 105]$
(%o23) $[0, 0, 0, 51, 19, 23, 767]$

(%o24)
$$\begin{pmatrix} 0 & 0 & 23 & 37 & 21 & 0 & 786 \\ 0 & 17 & 0 & 0 & 1 & 3 & 154 \\ 19 & 0 & 0 & 5 & 0 & 3 & 105 \\ 0 & 0 & 0 & 51 & 19 & 23 & 767 \end{pmatrix}$$

(%i102)

4 Application: le calcul du nombre de points entiers

Une application du caractère positif (au sens large) de la matrice étendue est la proposition suivante

Proposition 11. *Le nombre de points entiers d'un polytope*

Soit A=(a_{ij}) $\in \mathcal{M}_{mn}(\mathbb{N})$ et b=(b_i)$\in \mathcal{M}_{m1}(\mathbb{N})$ et le polytope P défini par le système $\begin{cases} AX = b \\ X \geqslant 0 \end{cases}$; le nombre de points entiers de P est le coefficient de $X^b = \prod_{i=1}^{m} x_i^{b_i}$ dans le développement en série entière de $\left(\prod_{j=1}^{n} \frac{1}{1 - X^{A_j}} \right)$, où $X^{A_j} = \prod_{i=1}^{m} x_i^{a_{ij}}$. [2]

Notons cependant la lourdeur de ce calcul (je n'ai pu l'exécuter sur mon PC dans le cas du polygone de l'exemple); comme la matrice étendue (A,b), à coefficients positifs, n'est nullement unique il pourrait être intéressant de chercher un moyen de la rendre « moins lourde », par exemple en réduisant les composantes du vecteur b.

Il serait intéressant aussi de comparer avec les calculs nécessaires pour appliquer la démarche de A. Barvinok.

Paris, Novembre 2019

Références:

[1] Alexandre Barvinok, Polynomial time algorithm for counting integral points in polyhedra when thedimension is fixed. Math. Oper. Res. 19, 769–779, 1994.

[2], Matthias Beck, The Partial-Fraction Method for counting integer Solutions of Integral Linear Systems, februray 2004, math.sfsu.edu/**beck**/papers/parfrac.pdf

[3] Shuzhong Zhang, New variants of finite Pivot Criss-Cross Algorithms for linear Programming, Report, 1999, http://www.menet.umn.edu/~zhangs/Reports/1999_Z.PDF.

[4] Komei Fukuda, Tomimo Matsui, On the Finiteness of th Criss-Cross Method, European Journal of Operations Research, Vol 52, 119-124,1991