C. R. Acad. Sci. Paris, Ser. I 335 (2002) 863–866

Combinatoire/*Combinatorics*

La valeur optimale des programmes entiers

Jean B. Lasserre

LAAS-CNRS, 7, avenue du Colonel Roche, 31077 Toulouse cedex 4, France

Reçu le 22 septembre 2002 ; accepté le 15 octobre 2002

Note présentée par Michèle Vergne.

Résumé On donne une expression de la valeur optimale $f_c(y)$ du programme entier $\max\{c'x \mid x \in \Omega(y) \cap \mathbb{N}^n\}$ où $\Omega(y)$ est le polyèdre convexe $\{x \in \mathbb{R}^n \mid Ax = y,\ x \geqslant 0\}$. Elle est une conséquence de la formule de Brion et Vergne qui évalue la somme $\sum_{x\in\Omega(y)\cap\mathbb{N}^n} \mathrm{e}^{c'x}$. On montre que comme en programmation linéaire, $f_c(y)$ peut être obtenue par inspection des coûts réduits aux sommets du polyèdre. On donne aussi un résultat explicite qui relie $f_c(ty)$ à la valeur optimale du programme linéaire associé, pour des valeurs de $t \in \mathbb{N}$ suffisamment grandes. ***Pour citer cet article : J.B. Lasserre, C. R. Acad. Sci. Paris, Ser. I 335 (2002) 863–866.***

The optimal value of integer programs

Abstract We present a formula for the optimal value $f_c(y)$ of the integer program $\max\{c'x \mid x \in \Omega(y) \cap \mathbb{N}^n\}$ where $\Omega(y)$ is the convex polyhedron $\{x \in \mathbb{R}^n \mid Ax = y,\ x \geqslant 0\}$. It is a consequence of Brion and Vergne's formula which evaluates the sum $\sum_{x\in\Omega(y)\cap\mathbb{N}^n} \mathrm{e}^{c'x}$. As in linear programming, $f_c(y)$ can be obtained by inspection of the reduced-costs at the vertices of the polyhedron. We also provide an explicit result that relates $f_c(ty)$ and the optimal value of the associated continous linear program, for large values of $t \in \mathbb{N}$. ***To cite this article: J.B. Lasserre, C. R. Acad. Sci. Paris, Ser. I 335 (2002) 863–866.***

1. Introduction

Etant donné $A \in \mathbb{Z}^{m\times n}$, $y \in \mathbb{Z}^m$, $c \in \mathbb{R}^n$, et le polyèdre convexe $\Omega(y) := \{x \in \mathbb{R}^n \mid Ax = y,\ x \geqslant 0\}$, on considère la fonction $f_c : \mathbb{Z}^m \to \mathbb{R}$

$$y \mapsto f_c(y) := \max\left\{c'x \mid x \in \Omega(y) \cap \mathbb{N}^n\right\} \tag{1}$$

dont la version *somme* $\widehat{f_c} : \mathbb{Z}^m \to \mathbb{R}$

$$y \mapsto \widehat{f_c}(y) := \sum\left\{\mathrm{e}^{c'x} \mid x \in \Omega(y) \cap \mathbb{N}^n\right\} \tag{2}$$

a été étudiée par plusieurs auteurs ces dernières années, notamment par Barvinok et Pommersheim [2], Brion et Vergne [3], Pukhlikov et Khovanskii [4]. En particulier, Brion et Vergne [3] donnent une formule explicite de $\widehat{f_c}(y)$ comme une somme pondérée sur les sommets de $\Omega(y)$, valide pour tout y dans une *chambre* γ. Dans cette note on donne une version « *max* » de la formule *somme* de Brion et Vergne, c'est-à-dire, une formule explicite de $f_c(y)$ pour tout y dans une chambre γ. On établit aussi un lien direct entre

Adresse e-mail : lasserre@laas.fr (J.B. Lasserre).

$f_c(ty)$ et la valeur optimale du programme linéaire associé $p_c^*(ty) := \max_{x\in\mathbb{R}^n}\{c'x \mid Ax = ty;\ x \geqslant 0\}$, pour y fixé et des valeurs de $t \in \mathbb{N}$ suffisamment grandes. Cette approche est une alternative aux approches *algébriques* décrites dans Thomas [6] et ses références.

2. Résultat principal

On introduit tout d'abord quelques définitions utilisées dans [3]. Soit $A \in \mathbb{Z}^{m\times n},\ c \in \mathbb{R}^n,\ y \in \mathbb{Z}^m$ avec $n > m$, et $\Omega(y) \subset \mathbb{R}^n$ le polyèdre convexe

$$\Omega(y) := \bigl\{x \in \mathbb{R}^n \mid Ax = y;\ x \geqslant 0\bigr\}. \tag{3}$$

Par la suite, c' et $c'x$ dénotent respectivement le transposé du vecteur c et le produit scalaire habituel de deux vecteurs c et x.

On note $A = [A_1|\dots|A_n]$ où $A_j \in \mathbb{Z}^m$ est la colonne j de A, pour tout $j = 1,\dots,n$. Avec $\Delta := (A_1,\dots,A_n)$ soit $C(\Delta) \subset \mathbb{R}^m$ le cône convexe fermé généré par la famille Δ. Un sous ensemble σ de $\{1,\dots,n\}$ est une *base* de Δ si la suite $\{A_j\}_{j\in\sigma}$ est une base de $\mathbb{R}^m$, et l'ensemble des bases de Δ est noté $\mathcal{B}(\Delta)$. Pour $\sigma \in \mathcal{B}(\Delta)$ soit $C(\sigma)$ le cône convexe généré par $\{A_j\}_{j\in\sigma}$. A tout $y \in C(\Delta)$ on associe l'intersection de tous les cônes $C(\sigma)$ qui contiennent y, ce qui définit une subdivision de $C(\Delta)$ en cônes polyédraux. Les interieurs des cônes *maximaux* de cette subdivision sont appelés *chambres* (cf. Alekseevskaya, Gel'fand et Zelevinsky [1]) avec fermeture notée $\overline{\gamma}$. Pour tout $y \in \gamma$, le polyèdre convexe $\Omega(y)$ en (3) est *simple*. Pour une chambre γ, soit $\mathcal{B}(\Delta,\gamma)$ l'ensemble des bases σ telles que $C(\sigma)$ contienne γ, et soit $\mu(\sigma)$ le volume du polytope $\{\sum_{j\in\sigma} t_jA_j \mid 0 \leqslant t_j \leqslant 1\}$. Pour tout $b \in \overline{\gamma}$ et $\sigma \in \mathcal{B}(\Delta,\gamma)$ on a $b = \sum_{j\in\sigma} x_j(\sigma)A_j$ pour des scalaires $x_j(\sigma) \geqslant 0$. Le vecteur $x(\sigma) \in \mathbb{R}^n_+$ avec $x_j(\sigma) = 0$ pour tout $j \notin \sigma$, est un *sommet* du polytope $\Omega(b)$. On note V le sous espace $\{x \in \mathbb{R}^n \mid Ax = 0\}$. Finalement, étant donné $\sigma \in \mathcal{B}(\Delta)$, soit $\pi^\sigma \in \mathbb{R}^m$ le vecteur ligne, solution unique de $\pi^\sigma A_j = c_j$ pour tout $j \in \sigma$. Le vecteur $c \in \mathbb{R}^n$ est dit *régulier* si $c_j - \pi^\sigma A_j \neq 0$ pour tout $\sigma \in \mathcal{B}(\Delta)$ et tout $j \notin \sigma$.

2.1. La formule « somme » discrète de Brion et Vergne

Soit $\widehat{f_c} : \mathbb{Z}^m \to \mathbb{R}_+$ la fonction définie en (2) où l'on suppose que $c \in \mathbb{R}^n$ est *régulier* et $-c$ est à l'intérieur du cône dual $(\mathbb{R}^n_+ \cap V)^*$. Soit $\Lambda := (A\mathbb{Z}^n) \subset \mathbb{Z}^m$. En utilisant la fonction génératrice de $\widehat{f_c}$, Brion et Vergne [3] (cf. extensions dans Szenes et Vergne [5]) montrent que pour tout $y \in \overline{\gamma} \cap \Lambda$

$$\widehat{f_c}(y) = \sum_{x(\sigma):\text{sommet de }\Omega(y)} \frac{\mathrm{e}^{c'x(\sigma)}}{\mu(\sigma)} U_\sigma(y,c), \tag{4}$$

où

$$U_\sigma(y,c) := \sum_{g\in G(\sigma)} \frac{\mathrm{e}^{2\mathrm{i}\pi y}(g)}{V_\sigma(g,c)} \quad \text{avec } V_\sigma(g,c) := \prod_{k\notin\sigma}\bigl(1 - \mathrm{e}^{-2\mathrm{i}\pi A_k}(g)\,\mathrm{e}^{c_k-\pi^\sigma A_k}\bigr), \tag{5}$$

et où $G(\sigma) := (\bigoplus_{j\in\sigma} \mathbb{Z}A_j)^*/\Lambda^*$ (où * dénote la dualité) est un groupe abélien fini d'ordre $\mu(\sigma)$, avec un nombre fini de caractères $\mathrm{e}^{2\mathrm{i}\pi y}$, $y \in \Lambda$; en particulier, en écrivant $A_k = \sum_{j\in\sigma} u_{jk}A_j$ pour tout $k \notin \sigma$, $\mathrm{e}^{2\mathrm{i}\pi A_k}(g) = \mathrm{e}^{2\mathrm{i}\pi\sum_{j\in\sigma} u_{jk}g_j}$. On remarque le rôle particulier de $c_k - \pi^\sigma A_k$, $k \notin \sigma$, qui est en programmation linéaire, le *coût réduit* de la variable *hors base* x_k par rapport à la base σ.

2.2. La version « max » de la formule de Brion et Vergne

On donne ici la version « max » de la formule discrète (ou *périodique*) (4) de Brion et Vergne [3]. On suppose d'abord que $c \in \mathbb{Q}^n$ et donc, après multiplication par un entier suffisamment grand, $c \in \mathbb{Z}^n$.

THÉORÈME 2.1. – *Soient $A \in \mathbb{Z}^{m\times n},\ b \in \mathbb{Z}^m,\ c \in \mathbb{Z}^n$, avec c régulier et $-c$ à l'intérieur du cone $(\mathbb{R}^n_+ \cap V)^*$. Soit $y \in \overline{\gamma} \cap \Lambda$ et soit q le plus petit commun multiple (p.p.c.m.) de $\{\mu(\sigma)\}_{\sigma\in\mathcal{B}(\Delta,\gamma)}$. Si $Ax = y$*

n'a pas de solution $x \in \mathbb{N}^n$, alors on pose $f_c(y) = -\infty$. Sinon, supposons que

$$\max_{\sigma \in \mathcal{B}(\Delta,\gamma)} \left[c'x(\sigma) + \lim_{r\to\infty} \frac{1}{r} \ln U_\sigma(y, rc) \right],$$

est atteint à un $\sigma \in \mathcal{B}(\Delta,\gamma)$ unique. Alors

$$f_c(y) = \max_{x(\sigma):\text{sommet de } \Omega(y)} \left[c'x(\sigma) + \lim_{r\to\infty} \frac{1}{r} \ln U_\sigma(y, rc) \right] \tag{6}$$

$$= \max_{x(\sigma):\text{sommet de } \Omega(y)} \left[c'x(\sigma) + \frac{1}{q} \big(\deg(P_{\sigma y}) - \deg(Q_{\sigma y})\big) \right] \tag{7}$$

pour certains polynômes réels $P_{\sigma y}, Q_{\sigma y}$ d'une variable. Le terme $\lim_{r\to\infty} \ln U_\sigma(y, rc)/r$ en (6) *ou* $(\deg(P_{\sigma y}) - \deg(Q_{\sigma y}))/q$ *en* (7)*, est la somme de certains coût-réduits $c_k - \pi^\sigma A_k$, $k \notin \sigma$.*

Démonstration. – On utilise la formule $\mathrm{e}^{f_c(y)} = \lim_{r\to\infty} (\widehat{f_{rc}}(y))^{1/r}$, ce qui donne

$$\mathrm{e}^{f_c(y)} = \lim_{r\to\infty} \big(\widehat{f_{rc}}(y)\big)^{1/r} = \lim_{r\to\infty} \left[\sum_{x(\sigma):\text{sommet de } \Omega(y)} \frac{\mathrm{e}^{rc'x(\sigma)}}{\mu(\sigma)} \sum_{g\in G(\sigma)} \frac{\mathrm{e}^{2\mathrm{i}\pi y}(g)}{V_\sigma(g, rc)} \right]^{1/r}$$

$$= \lim_{r\to\infty} \left[\sum_{x(\sigma):\text{sommet de } \Omega(y)} H_\sigma(y, rc) \right]^{1/r}. \tag{8}$$

Avec rc au lieu de c en (5), on voit que $V_\sigma(g, rc)$ est une fonction de la variable $v := \mathrm{e}^r$ et donc, à son tour, $H_\sigma(y, rc)$ est une fonction de v de la forme

$$H_\sigma(y, rc) = \big(\mathrm{e}^r\big)^{c'x(\sigma)} \sum_{g\in G(\sigma)} \frac{\mathrm{e}^{2\mathrm{i}\pi y}(g)}{\sum_j (\delta_j(\sigma, g, A) \times (e^r)^{\alpha_j(\sigma,c)})}, \tag{9}$$

pour certains coefficients $\{\delta_j(\sigma, g, A), \alpha_j(\sigma, c)\}$. Les coefficients $\alpha_j(\sigma, c)$ sont des sommes de coûts réduits $c_k - \pi^\sigma A_k$, $k \notin \sigma$, qui ne *dépendent pas* de y, comme d'ailleurs les coefficients (complexes) $\{\delta_j(\sigma, g, A)\}$.

Soit $v := \mathrm{e}^{r/q}$ où q est le p.p.c.m. de $\{\mu(\sigma)\}_{\sigma\in\mathcal{B}(\Delta,\gamma)}$. Comme $q(c_k - \pi^\sigma A_k) \in \mathbb{Z}$ pour tout $k \notin \sigma$,

$$H_\sigma(y, rc) = v^{qc'x(\sigma)} \times \frac{P_{\sigma y}(v)}{Q_{\sigma y}(v)}, \tag{10}$$

pour des polynômes réels $P_{\sigma y}, Q_{\sigma y} \in \mathbb{R}[v]$. Donc quand $r \to \infty$,

$$H_\sigma(y, rc) \approx \big(\mathrm{e}^{r/q}\big)^{qc'x(\sigma)+\deg(P_{\sigma y})-\deg(Q_{\sigma y})}, \tag{11}$$

de telle sorte que la limite en (8) est donnée par

$$\max_{x(\sigma):\text{sommet de } \Omega(y)} \mathrm{e}^{c'x(\sigma)+(\deg(P_{\sigma y})-\deg(Q_{\sigma y}))/q},$$

ce qui donne (6), (7). □

Comme pour le programme linéaire associé $p_c^*(y) := \max_{x\in\mathbb{R}^n}\{c'x \mid Ax = y,\ x \geqslant 0\}$, la valeur optimale $f_c(y)$ peut être obtenue par inspection (mais plus compliquée, via (7)) des *coûts réduits* aux sommets.

En considérant maintenant le polytope dilaté $\Omega(ty)$, $t \in \mathbb{N}$, on obtient le résultat suivant.

COROLLAIRE 2.2. – *Soient $A \in \mathbb{Z}^{m\times n}$, $b \in \mathbb{Z}^m$, $c \in \mathbb{Z}^n$, avec c régulier et $-c$ à l'intérieur du cone $(\mathbb{R}^n_+ \cap V)^*$. Soit $y \in \gamma \cap \Lambda$ et soit $x(\sigma^*) \in \Omega(y)$ un sommet optimal du programme linéaire $\mathrm{p}_c^*(y) := \max_{x\in\mathbb{R}^n}\{c'x \mid Ax = y,\ x \geqslant 0\}$, i.e., $p_c^*(y) = c'x(\sigma^*)$, $\sigma^* \in \mathcal{B}(\Delta,\gamma)$. Alors, pour $t \in \mathbb{N}$ suffisamment grand*

$$f_c(ty) - p_c^*(ty) = \lim_{r\to\infty} \left[\frac{1}{r} \ln U_{\sigma^*}(ty, rc) \right]. \tag{12}$$

En particulier, pour $t \in \mathbb{N}$ suffisamment grand, la fonction $t \mapsto f_c(ty) - p_c^*(ty)$ est périodique (constante) de période $\mu(\sigma^*)$, et est une somme de coûts réduits par rapport à la base optimale $\sigma^* \in \mathcal{B}(\Delta, \gamma)$.

Démonstration. – Soit $t \in \mathbb{N}$ et notons que $p_{rc}^*(ty) = \operatorname{tr} p_c^*(y) = \operatorname{tr} c'x(\sigma^*)$. Comme dans la preuve du Théorème 2.1, et avec ty au lieu de y, on a

$$\widehat{f}_{rc}(ty)^{1/r} = \mathrm{e}^{tc'x(\sigma^*)}\left[\frac{U_{\sigma^*}(ty, rc)}{\mu(\sigma^*)} + \sum_{\text{sommet } x(\sigma)\neq x(\sigma^*)} \left(\frac{\mathrm{e}^{rc'x(\sigma)}}{\mathrm{e}^{rc'x(\sigma^*)}}\right)^t \frac{U_\sigma(ty, rc)}{\mu(\sigma)}\right]^{1/r}$$

et au vu de (9), (10), posant $\delta_\sigma := c'x(\sigma^*) - c'x(\sigma) \geqslant 0$ et $v := \mathrm{e}^{r/q}$,

$$\widehat{f}_{rc}(ty)^{1/r} = \mathrm{e}^{tc'x(\sigma^*)}\left[\frac{U_{\sigma^*}(ty, rc)}{\mu(\sigma^*)} + \sum_{\text{sommet } x(\sigma)\neq x(\sigma^*)} v^{-tq\delta_\sigma} \frac{P_{\sigma ty}(v)}{Q_{\sigma ty}(v)}\right]^{1/r}.$$

En fait, $\delta_\sigma > 0$ si $\sigma \neq \sigma^*$ parce que $\Omega(y)$ est simple si $y \in \gamma$, et c est régulier. Donc, si $x(\sigma^*)$ est un sommet optimal du programme linéaire associé, les coûts réduits $c_k - \pi^{\sigma^*} A_k$, $k \notin \sigma^*$, par rapport à la base optimale σ^*, sont alors tous non positifs, et en fait, strictement négatifs si c est régulier. Alors, $c'x(\sigma^*) > c'x(\sigma)$ si $\sigma \neq \sigma^*$. Donc, au vu du coefficient $\mathrm{e}^{-tq\delta_\sigma}$ et du nombre fini de sommets, quand $r \to \infty$, le terme

$$\sum_{\text{sommet } x(\sigma)\neq x^*} v^{-tq\delta_\sigma} \frac{P_{\sigma ty}(v)}{Q_{\sigma ty}(v)}$$

est négligeable devant $U_{\sigma^*}(ty, rc)$ pour t suffisamment grand, parce que les degrés et coefficients de $P_{\sigma ty}$ and $Q_{\sigma ty}$ qui dépendent de ty, mais pas de la grandeur de ty (cf. (9), (10)) sont tous bornés en ty. Le passage à la limite quand $r \to \infty$, donne

$$\mathrm{e}^{f_c(ty)} = \mathrm{e}^{tc'x(\sigma^*)} \lim_{r\to\infty} U_{\sigma^*}(ty, rc)^{1/r},$$

d'où l'on déduit (12). Finalement, la périodicité vient du terme $\mathrm{e}^{2\mathrm{i}\pi ty}(g)$ en (5) pour $g \in G(\sigma^*)$. La période est donc l'ordre de $G(\sigma^*)$, i.e., le volume du polytope $\{\sum_{j\in\sigma^*} t_j A_j \mid 0 \leqslant t_j \leqslant 1\}$ (ou le déterminant de la matrice $[A_j]_{j\in\sigma^*}$). $\square$

Théoriquement, on peut obtenir les polynômes $P_{\sigma y}$, $Q_{\sigma y}$ par le calcul symbolique. Finalement si $c \in \mathbb{R}^n$ les formules (6), (7) et (12) sont aussi valides mais les fonctions $P_{\sigma y}$, $Q_{\sigma y}$ ne sont plus des polynômes.

Références bibliographiques

[1] T.V. Alekseevskaya, I.M. Gel'fand, A.V. Zelevinskij, An arrangement of real hyperplanes and the partition function connected with it, Soviet Math. Dokl. 36 (1988) 589–593.

[2] A.I. Barvinok, J.E. Pommersheim, An algorithmic theory of lattice points in polyhedra, in: New Perspectives in Algebraic Combinatorics, in: MSRI Publications, Vol. 38, 1999, pp. 91–147.

[3] M. Brion, M. Vergne, Residue formulae, vector partition functions and lattice points in rational polytopes, J. Amer. Math. Soc. 10 (1997) 797–833.

[4] A.V. Pukhlikov, A.G. Khovanskii, A Riemann–Roch theorem for integrals and sums of quasipolynomials over virtual polytopes, St. Petersburg Math. J. 4 (1993) 789–812.

[5] A. Szenes, M. Vergne, Residue formulae for vector partitions and Euler–Maclaurin sums, Adv. Appl. Math., à paraître.

[6] R. Thomas, Algebraic methods in integer programming, in: C. Floudas, P. Pardalos (Eds.), Encyclopedia of Optimization, Kluwer Academic, Dordrecht, 2001.