Article invité

L'Approche par Programmation Semidéfinie en Optimisation Combinatoire

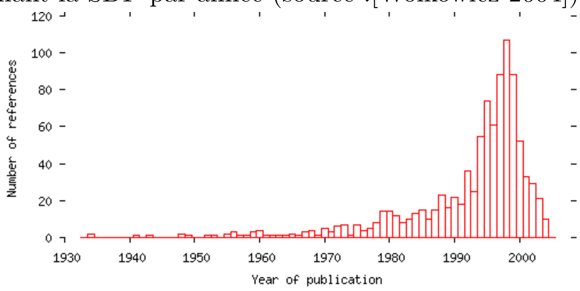
Frédéric Roupin[1]
roupin@iie.cnam.fr

Résumé

La programmation semidéfinie est connue pour les avancées qu'elle a rendues possibles en approximation au pire cas de problèmes difficiles de l'optimisation combinatoire. Elle est également réputée comme une approche coûteuse en temps de calcul, et donc difficilement exploitable dans la pratique. Qu'en est-il réellement ? Cet article replace dans leur contexte un ensemble de références bibliographiques permettant de mieux appréhender ce domaine de recherche encore très jeune et actif.

1 Introduction

La programmation semidéfinie (SDP) a provoqué un réel engouement dans la communauté de l'optimisation combinatoire après les fameux résultats de Goemans et Williamson concernant l'approximation du problème MAX-CUT [Goemans 1995]. Une série d'articles connexes est alors parue (par exemple [Karger at al 1994, Alon et Kahale 1995, Freize et Jerrum 1995]). Cette "explosion" est très visible sur la figure suivante qui donne le nombre de publications concernant la SDP par année (source :[Wolkowicz 2004]).



Ces résultats concernant l'approximation au pire cas, bien que remarquables, limitaient l'utilisation de la SDP au plan théorique car la résolution numérique des SDP restait très coûteuse en temps. D'autre part, il manquait des méthodes générales et systématiques pour l'élaboration de relaxations semidéfinies, i.e. ne relevant pas d'une simple application "d'astuces". En effet, les relaxations semidéfinies proposées étaient la plupart du temps conçues "ex nihilo" en suivant une recette à la "Goemans-Williamson". Enfin, un goût certain pour l'ésotérisme dans quelques notations, et l'utilisation courante de variables bivalentes à valeurs dans {-1,1} pour établir les modèles utilisés pour obtenir les relaxations SDP ne facilitaient ni la lecture ni la comparaison des différentes approches, pour la plupart des néophytes. Ce contexte représentait un obstacle certain pour la diffusion et l'utilisation pratique de cette approche (pour la résolution exacte ou approchée).

Parallèlement à tous ces résultats théoriques sur l'approximation de problèmes difficiles, des progrès fulgurants furent accomplis dans le domaine des algorithmes de points intérieurs. Ainsi, les travaux de [Nesterov 1994, Alizadeh 1995] furent les premiers représentants de méthodes de plus en plus efficaces (voir Section 3). Grâce à ces nouveaux outils, plusieurs expérimentations numériques furent menées (par exemple [Zhao 1998]), conduisant à de meilleures bornes mais également à des méthodes exactes de résolution (par exemple [Cung Roupin 1999]). Plus récemment, des études ont comparé les différentes approches (linéaire, lagrangienne, semidéfinie) et les liens entre elles ont été clarifiés (voir Section 4). Enfin en revisitant les méthodes utilisées en programmation linéaire et en profitant de l'effet de "levier" de la contrainte (non linéaire) de positivité, plusieurs schémas ou algorithmes ont été élaborés pour construire des relaxations semidéfinies (voir Section 5).

2 Qu'est-ce qu'un programme semidéfini ?

Le produit scalaire habituel sur S_n (espace des matrices symétriques réelles $n \times n$) est

$$(A, B) \in S_n^2 \;\rightarrow\; A \bullet B = Tr\,(AB) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ij}$$

[1]CEDRIC-Institut d'Informatique d'Entreprise, 18 allée Jean Rostand 91025 Evry cedex, France

On peut ainsi réécrire toute forme quadratique $x^T Ax+b^T x+c$ comme $A \bullet xx^T+b^T x+c$, où $x \in \mathbb{R}^n$. On considère $S_n^+ = \{A : \forall z \in \mathbb{R}^n, z^T Az \geq 0\} \subset S^n$ l'ensemble des matrices (semidéfinies) positives. Pour $A \in S_n^+$ on écrit également $A \succeq 0$. Un programme semidéfini peut alors être défini comme la maximisation d'une fonction linéaire de $X \in S_n^+$ soumise à des contraintes linéaires :

$$(SDP) \left\{ \begin{array}{lll} \text{Max} & A_0 \bullet X & \\ \text{s.c.} & A_i \bullet X = c_i & i = 1, ..., m \\ & X \succeq 0 & \end{array} \right.$$

où $c \in \mathbb{R}^m$, et $A_i \in S_n$ $\forall i \in \{0, ..., m\}$. Le programme dual de (SDP) est :

$$(DSDP) \left\{ \begin{array}{ll} \text{Min} & c^T y \\ \text{s.c.} & F(y) = \sum_{i=1}^m A_i y_i - A_0 \succeq 0 \\ & y \in \mathbb{R}^m \end{array} \right.$$

En fait, $(DSDP)$ est également un programme semidéfini. Pour s'en convaincre, il suffit de constater que les matrices s'écrivant $F(y) = \sum_{i=1}^m A_i y_i - A_0$ sont éléments de l'espace affine de repère $(A_0, ..., A_m)$ (la matrice $-A_0$ est donc l'origine, et les y_i les coordonnées de $F(y)$ dans la base). En décrivant cet espace par un ensemble de contraintes linéaires, on peut reformuler $(DSDP)$ comme (SDP) (avec des matrices A_i et des vecteurs c_i différents bien sûr). La programmation semidéfinie peut être vue comme une généralisation de la programmation linéaire, puisque si toutes les matrices A_i pour $i \in \{0, ..., m\}$ sont diagonales alors (SDP) et $(DSDP)$ deviennent de simples programmes linéaires. On peut également vérifier que la programmation quadratique convexe est un cas particulier de la programmation semidéfinie.

3 Résoudre un programme semidéfini

Un des problèmes majeurs rencontrés lorsque l'on souhaite utiliser la programmation semidéfinie est l'absence d'outils numériques commerciaux éprouvés. Cela est dû en partie au fait que l'élaboration de méthodes de résolution de SDP est un domaine de recherche récent et encore très actif. Il existe cependant à présent un certain nombre d'outils gratuits, mais qui réclament une certaine habitude et dont l'efficacité est variable suivant les problèmes traités [Borchers 1999, Sturm 1999, Helmberg 2000d, Burer Monteiro 2003, Tütüncü et al 2003].

Les premiers outils de résolution mettaient près d'une heure pour venir à bout d'une relaxation semidéfinie d'un MAXCUT d'un graphe de cinquante sommets. A présent, grâce à l'outil SB ("Spectral Bundle method") [Helmberg 2000d], il est possible d'obtenir la valeur optimale du problème "had14" (QAP : affectation quadratique) en résolvant en moins de dix minutes sur un simple PC une relaxation semidéfinie [Roupin 2004] impliquant 19701 variables et 24851 contraintes. Des programmes comportant plus de 400000 variables et 450000 contraintes peuvent être traités en moins d'une journée (par exemple une relaxation semidéfinie de nug30, une instance bien connue d'affectation quadratique). SB est l'implémentation d'une méthode de sous-gradient [Helmberg 2000c] qui résout le dual de SDP dont la trace de la matrice est constante (et donc en fait bornée en introduisant une variable d'écart). Si la borne est très grande, la convergence peut être très lente et d'autres outils (comme CSDP, une méthode primal-dual) sont beaucoup plus efficaces.

De nombreuses expérimentations ont par conséquent été menées [Helmberg 2000a, Benson 2000, Anjos 2002, Helmberg 2003, Roupin 2004], et on peut trouver des études comparatives des différentes méthodes et logiciels existants (par exemple [Mittelmann 2003]).

4 Approches lagrangienne et Semidéfinie

Le lien entre relaxations lagrangienne et semidéfinie est clairement établi dans [Poljak 1995]. D'autre part, dans [Lemarechal 1999] il est démontré (entre autres) que les relaxations semidéfinies proposées dans [Goemans 1995] pour le problème MAX-CUT peuvent être obtenues en dualisant n contraintes $x_i^2 = 1$, une formulation naturelle des contraintes de bivalence $x_i \in \{-1, 1\}$ présentes dans un modèle du problème. Dans ce même article, les auteurs établissent le lien entre relaxation lagrangienne partielle d'un programme quadratique en 0-1 (les contraintes linéaires sont maintenues) et programmes semidéfinis. Ce lien peut être étendu [Faye Roupin 2004b] au cas plus général d'un problème quadratique quelconque (x n'étant donc pas nécessairement booléen) : le dual d'une relaxation lagrangienne partielle (les contraintes linéaires sont maintenues) peut être également formulé comme un programme semidéfini avec cependant plus de contraintes que dans le cas booléen. Dans ce dernier cas, les deux SDP sont équivalents, mais il existe un point admissible de la relaxation lagrangienne correspondant à la relaxation semidéfinie présentée dans

[Faye Roupin 2004b] qui atteint la valeur optimale. Ceci influence fortement les temps de résolution des programmes semidéfinis correspondants. Une meilleure compréhension des liens entre les deux approches a donc des conséquences directes sur le plan pratique.

5 Construire une relaxation semidéfinie

L'article fondateur de Goemans-Williamson a influencé de nombreux auteurs. L'idée de base de cette approche est de remplacer chacune des n variables (à valeur dans $\{-1, 1\}$) du modèle discret initial par un vecteur unitaire v_i. L'interprétation géométrique de la matrice $Y = V^T V$ (variable du SDP) comme la donnée de ce champ de vecteurs $V = [v_1, ..., v_n]$ de la sphère unité a conduit à l'élaboration de plusieurs algorithmes avec garanties de performance. La relaxation SDP peut également être vue comme une relaxation sur le rang de la matrice $Y = V^T V$ (qui vaut 1 avant relaxation). Le problème technique de la representation des termes linéaires dans ce SDP est simplement résolu en ajoutant un vecteur de référence v_0 (uniquement contraint à être de norme 1), et donc une ligne et une colonne à la matrice Y (les produits scalaires $v_0^T v_i$). Bien que fertile pour construire des algorithmes approchés, cette approche est en fait équivalente à celle proposée dans [Shor 1987, Lovasz 1991] puis dans [Poljak 1995], où un modèle plus "classique" en variables booléennes (à valeurs dans $\{0, 1\}$) est défini. En effet, les relaxations semidéfinies issues de ces deux approches sont équivalentes [Laurent 1997]. Le passage de l'une à l'autre étant simplement obtenu en utilisant l'automorphisme $X \rightarrow QXQ^T$ de S_n^+, où $Q = \begin{bmatrix} 1 & 0 \\ \frac{1}{2}e_n & \frac{1}{2}I_n \end{bmatrix}$ (e_n est le vecteur dont tous les éléments valent 1) [Helmberg 2000b]. On peut en particulier établir les formulations correspondantes des contraintes de linéarisation standard dans les deux modèles (i,j et k sont dans $\{1, ..., n\}$) :

modèle $\{-1, 1\}$	modèle $\{0, 1\}$
$Y_{0i} + Y_{0j} + Y_{ij} \geq -1$	$X_{ij} \geq 0$
$Y_{ij} - Y_{i0} - Y_{j0} \geq -1$	$x_i + x_j \leq 1 + X_{ij}$
$Y_{i0} - Y_{j0} - Y_{ij} \geq -1$	$X_{ij} \leq x_i$
$-Y_{i0} + Y_{j0} - Y_{ij} \geq -1$	$X_{ij} \leq x_j$

Plusieurs "recettes" ont été proposées pour améliorer ces relaxations semidéfinies [Lemarechal 1999, Helmberg 2000b, Lasserre 2002a], ainsi que des comparaisons entre les différents schémas existants [Lasserre 2000b, Laurent 2003]. D'autre part, dans [Roupin 2004], un algorithme est proposé pour construire des relaxations semidéfinies à partir de n'importe quelle relaxation linéaire. Les avantages principaux de cette démarche sont de garantir l'obtention d'une relaxation SDP plus efficace que la relaxation PL servant à la construire, et de profiter de toute l'expérience accumulée dans le domaine de la programmation linéaire afin d'identifier des coupes efficaces. Des résultats théoriques et numériques sont présentés pour les problèmes suivants : l'affectation quadratique (QAP), la recherche d'un sous-graphe dense (k-cluster), et un problème de placement de tâches dans un système distribué avec contraintes de ressources. Enfin, grâce au caractère algorithmique de cette démarche, un modeleur automatique a été développé (*SDP_S*), et est disponible sur le site *http : //semidef.free.fr* [SDPS 2003], la résolution numérique des programmes semidéfinis obtenus étant prise en charge par SB [Helmberg 2000d].

6 Coupes et approche semidéfinie

Un autre domaine très actif actuellement est l'utilisation de coupes pour améliorer les bornes mais également les temps de résolution des relaxations semidéfinies existantes. Ainsi, plusieurs articles présentent des résultats d'expérimentations numériques issus de cette démarche. Par exemple, dans [Helmberg, 2001] un algorithme de coupe utilisant les inégalités de cycles de longueur impaire est décrit pour les problèmes MAXCUT, EQUICUT et celui de la bisection d'un graphe. Dans [Faye Roupin 2004a], nous avons amélioré les bornes obtenues dans [Roupin 2004] (ainsi que diminué les temps de calcul de 60 à 80%), en ajoutant à une relaxation semidéfinie du QAP (affectation quadratique) des coupes appartenant aux familles présentées dans [Blanchard et al 2000]. Il existe également un article de synthèse sur les différentes tentatives SDP+coupes [Mitchell 2002]. Des coupes non linéaires (en fait quadratiques convexes) ont été présentées dans [Iyengar, 2001b], et des techniques généralisant celles de "lift-and-project" de la programmation linéaire [Adams 1990, Lovasz 1991, Balas 1993] ont été explorées dans [Iyengar 2001a] (toujours avec l'objectif d'engendrer des coupes pour des SDP).

7 Conclusion

D'un point de vue théorique, la SDP a permis d'améliorer significativement les résultats sur l'approximation de nombreux problèmes. Sur le plan pratique, il est à présent possible de résoudre des programmes semidéfinis de taille moyenne et même grande pour des instances suffisamment "creuses" (voir section 3). Cependant, il est clair que l'approche semidéfinie doit être réservée au traitement de problèmes très difficiles, en particulier pour ceux où la programmation linéaire est inefficace. En effet, le gain obtenu sur la borne doit compenser la perte de temps provoquée par la résolution d'un SDP. Cependant, les derniers progrès des outils numériques et l'utilisation de relaxations SDP de plus en plus élaborées montrent que cette approche est dorénavant tout à fait exploitable dans la pratique, et ceci même dans le cadre de la résolution exacte de certains problèmes.

Références

[Adams 1990] W.P. Adams et H.D. Sherali, "A hierarchy of relaxations between the continuous and convex hull representations of zero-one programming problems", *SIAM J. Disc. Math* 3, pp. 411-430, 1990.

[Alizadeh 1995] F. Alizadeh "Interior point methods in semidefinite programming with applications to combinatorial optimization", *SIAM J. Optim.*, 5(1), pp. 13-51, 1995.

[Alon et Kahale 1995] N. Alon et N. Kahale, "Approximating the independence number via the theta-function" Technical report, Tel Aviv University, Israel.

[Anjos 2002] M. F. Anjos. "An Improved Semidefinite Programming Relaxation for the Satisfiability Problem", Technical Report UW-E&CE#2002-09, University of Waterloo, Canada, June 2002.

[Balas 1993] E. Balas, S. Ceria, G. Cornuejols, "A lift-and-project cutting plane algorithm for mixed 0-1 programs". *Math. Prog. A* 58 pp. 295-323, 1993.

[Benson 2000] S. Benson, Y. Ye, X. Zhang, "Solving large-scale sparse semidefinite programs for combinatorial optimization" *SIAM J. Optim.*, 10(2), pp. 443-461, 2000.

[Blanchard et al 2000] A. Blanchard, S. Elloumi, A. Faye, N. Wicker. "Un algorithme de coupes pour l'Affectation Quadratique". In Proceedings FRANCORO III, Québec, Canada Mai 2001.

[Burer Monteiro 2003] S. Burer, R.D. Monteiro, "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization." *Math Prog.* 95(2), pp 329-357, 2003.

[Borchers 1999] Brian Borchers. CSDP, A C Library for Semidefinite Programming. *Optimization Methods and Software* 11(1) :613-623, 1999. *http ://www.nmt.edu/ borchers/csdp.html*

[Cung Roupin 1999] Van Dat Cung et Frédéric Roupin, "A Parallel Branch-and-Bound Algorithm using a Semidefinite Programming Relaxation for the Vertex-Cover Problem" In Proceedings of ECCO XII, Bandol, France, 1999.

[Faye Roupin 2004a] A.Faye et F.Roupin, "A lower bound for the Quadratic Assignment Problem based upon a semidefinite relaxation and a cutting planes approach" In Proceedings ECCO 2004, June 24-26, Beirut, Lebanon, 2004.

[Faye Roupin 2004b] A. Faye, F. Roupin, "Partial Lagrangian and Semidefinite relaxations of Quadratic Programs". Rapport technique CEDRIC No 673, 2004. http ://cedric.cnam.fr.

[Freize et Jerrum 1995] A. Frieze et M. Jerrum. Improved approximation algorithms for MAX k-CUT and MAX BISECTION. IPCO IV Proc., LNCS 920, Springer 1995, pp. 1-13.

[Goemans 1995] M.X. Goemans, S.P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming" *Journal of the ACM* 42(6) :1115-1145, 1995.

[Helmberg 1998] C. Helmberg, F. Rendl. Solving quadratic (0,1)-problems by semidefinite programs and cutting planes" *Math. Programming*, 82(3, Serie A), pp. 291-315, 1998.

[Helmberg 2000a] C. Helmberg, F. Rendl, et R. Weismantel, "A Semidefinite Programming Approach to the Quadratic Knapsack Problem" *J. of Comb. Opt.* Vol. 4 pp. 197-215, 2000.

[Helmberg 2000b] C. Helmberg, "Semidefinite Programming for Combinatorial Optimization, Habilitationsschrift" ZIB-report ZR-00-34, KZZI, Takustrasee 7, 14195 Berlin, Germany, 2000.

[Helmberg 2000c] C. Helmberg, F. Rendl, "A spectral bundle method for semidefinite programming" *SIAM J. Optim.* 10(3) :673-696, 2000.

[Helmberg 2000d] C. Hemberg, "A C++ implementation of the Spectral Bundle Method", Manual version 1.1.1. *http ://www-user.tu-chemnitz.de/ helmberg/SBmethod/.*

[Helmberg, 2001] C. Helmberg, "Cutting planes algorithm for large scale semidefinite relaxations".

ZIB-Report ZR 01-26, KZZI, Takustraβe 7, 14195 Berlin, Germany, 2001.

[Helmberg 2003] C. Helmberg, “Numerical Evaluation of SBmethod”, *Math Prog.* 95(2), pp 381-406, 2003.

[Iyengar 2001a] G. Iyengar et M. T. Cezik, “Cutting planes for mixed 0-1 mixed semidefinite programs.” Integer and Combinatorial Optimization, LNCS 2081, pp. 251-263, 2001.

[Iyengar, 2001b] G. Iyengar, “Quadratic cuts for mixed 0-1 quadratic programs”, Technical report, IEOR Dept., Columbia University, 2001.

[Karger at al 1994] D. Karger, R. Motwani, et M. Sudan. Approximate Graph Coloring by Semidefinite Programming. FOCS'94, pp. 2-13, 1994.

[Lasserre 2002a] J.B. Lasserre, ”An Explicit Exact SDP Relaxation for Nonlinear 0-1 Programs” *SIAM J. Optim.* 12, pp. 756-769 2002.

[Lasserre 2000b] J.B. Lasserre, ”Semidefinite programming vs. LP relaxations for polynomial programming” *Math. Oper. Res.* 27(2) :347-360, 2002.

[Laurent 1997] M. Laurent, S.Poljak, et F. Rendl, “Connections between semidefinite relaxations of the max-cut and stable set Problems” *Mathematical Programming* 77 pp. 225-246, 1997.

[Laurent 2003] M. Laurent, “A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming” *Mathematics of Oper. Res.* 28, pp. 470-496, 2003.

[Lemarechal 1999] C. Lemarechal et F. Oustry, “Semidefinite relaxations and Lagrangian duality with application to combinatorial optimization”, RR-3710, INRIA Rhone-Alpes, ZIRST - 655 avenue de l'Europe, F-38330 Montbonnot Saint-Martin, June 1999.

[Lovasz 1991] L. Lovász, A. Schrijver, ”Cones of matrices and set functions and 0-1 optimization”, *SIAM J. on Opt.* vol 1 :166-190, 1991.

[Mitchell 2002] J. E. Mitchell , K.Krishnan , “Cutting plane methods for semidefinite programming”, Technical report, Rensselaer Polytechnic Institute Troy, NY 12180 USA, 2002.

[Mittelmann 2003] Hans D. Mittelmann, ”An Independent Benchmarking of SDP and SOCP Solvers” *Math Prog.* 95(2), pp. 407-430, 2003. *http ://plato.asu.edu/bench.html.*

[Nesterov 1994] Y. Nesterov, A. Nemirovski, ”Interior-Point Polynomial Algorithms in Convex Programming” *SIAM Studies in Applied Mathematics* Philadelphia, 1994.

[Poljak 1995] S. Poljak, F.Rendl, H. Wolkowicz, “A recipe for semidefinite relaxation for (0,1)-quadratic programming” *Journal of Global Optimization* Vol. 7, pp. 51-73, 1995.

[Roupin 2004] F. Roupin, ”From Linear to Semidefinite Programming : an Algorithm to obtain Semidefinite Relaxations for Bivalent Quadratic Problems” *Journal of Combinatorial Optimization* vol. 8(4), pp. 469-493, 2004.

[SDPS 2003] G. Delaporte, S. Jouteau, F. Roupin, “SDP_S : a Tool to formulate and solve Semidefinite relaxations for Bivalent Quadratic problems”, In Proceedings ROADEF 2003, Avignon 26-28 Février, 2003. *http ://semidef.free.fr.*

[Shor 1987] N.Z. Shor, “Quadratic optimization problems.” *Soviet Journal of computer Systems Sciences* 25, pp 1-11, 1987.

[Sturm 1999] J.F. Sturm, “Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones” *Optimization Methods and Software* 11, pp. 625-653, 1999.

[Tütüncü et al 2003] Reha H. Tütüncü, K. C. Toh, Michael J. Todd, ”Solving semidefinite-quadratic-linear programs using SDPT3” *Math Prog.* Vol 95(2) pp 189-217, 2003.

[Wolkowicz 2004] Henry Wolkowicz, ”A Bibliography on Semidefinite Programming” *http ://liinwww.ira.uka.de/bibliography/Math/.*

[Zhao 1998] Q. Zhao, S.E. Karish, F. Rendl, et H. Wolkowicz. “Semidefinite programming relaxations for the quadratic assignment problem”. *J. of Comb. Opt.* 2(1) :71-109, 1998.