Méthodes Matricielles
Introduction à la Complexité Algébrique

ABDELJAOUED Jounaïdi
Maître Assistant
Université de Tunis

LOMBARDI Henri
Maître de Conférences
Université de Franche-Comté, Besançon

2003

Avant-Propos

L'algèbre linéaire nous a semblé être le terrain idéal pour une introduction simple et pédagogique aux outils modernes de la complexité algébrique développés durant les trois dernières décennies.

Le tournant en matière de complexité algébrique en algèbre linéaire fut la découverte par Strassen [86], en 1969, d'un fait d'une simplicité déconcertante, mais d'une portée considérable, à savoir que la multiplication de deux matrices carrées d'ordre deux pouvait se faire avec seulement sept multiplications (non commutatives) au lieu de huit dans l'anneau de base. Ce qui ramenait la complexité asymptotique de la multiplication de deux matrices carrées d'ordre n à $\mathcal{O}(n^{\log_2 7})$ au lieu de $\mathcal{O}(n^3)$ et faisait descendre pour la première fois l'exposant de n au-dessous de 3, alors que les recherches antérieures n'avaient réussi qu'à réduire le coefficient de n^3 dans le nombre d'opérations arithmétiques nécessaires pour calculer le produit de deux matrices carrées d'ordre n (cf. [18]).

Depuis, de nombreux outils ont été développés. Des notions nouvelles sont apparues comme celles de complexité bilinéaire et de rang tensoriel utilisées de manière intensive notamment par Bini, Pan, Schönhage, Strassen, Winograd et d'autres (cf. [7, 8, 19, Pan, 82, 90]) pour réduire l'exposant α : à l'heure actuelle, on sait que $\alpha < 2,376$. Il est cependant conjecturé que la borne inférieure des exposants α acceptables serait 2, c'est-à-dire que pour tout $\varepsilon > 0$ le produit de deux matrices carrées d'ordre n pourrait être calculé par un circuit arithmétique de taille $\mathcal{O}(n^{2+\varepsilon})$ et de profondeur $\mathcal{O}(\log n)$. Cependant ces méthodes, d'un intérêt théorique certain, sont à l'heure actuelle inapplicables à cause notamment de la constante démesurée que le « *grand* $\mathcal{O}$ » cache (cf. [Knu] § 4.6.4). Par contre la méthode de Strassen a pu trouver une implémentation concrète [14], et elle commence à battre la multiplication usuelle (dite conventionnelle) à partir de $n = 70$.

Le calcul parallèle est une technique, en plein développement, qui distribue un calcul à faire sur un grand nombre de processeurs travaillant au même moment, en parallèle. Pour la multiplication rapide de matrices carrées, si le nombre de processeurs disponibles est suffisamment grand (de l'ordre de $\mathcal{O}(n^\alpha)$), le temps de calcul est alors extrêmement faible (de l'ordre de $\mathcal{O}(\log n)$ pour des matrices sur un corps fini).

La multiplication rapide des matrices carrées a de nombreuses applications en algèbre linéaire sur les corps, par exemple l'inversion d'une matrice carrée peut se faire en $\mathcal{O}(n^\alpha)$ avec le même exposant. Cependant, contrairement à la multiplication rapide des matrices, ces algorithmes ne sont pas bien adaptés au *calcul parallèle*. Ainsi l'agorithme d'inversion d'une matrice carrée auquel on vient de faire allusion, et que nous étudierons dans la section 8.2, ne voit jamais son temps de calcul descendre en dessous d'un $\mathcal{O}(n \log n)$.

C'est sur la base de résultats parfois anciens qu'on a pu exhiber, en algèbre linéaire, des algorithmes bien adaptés au calcul parallèle, s'appuyant sur la multiplication rapide des matrices. Ces algorithmes sont en outre des algorithmes sans divisions (ou presque) et s'appliquent donc à des anneaux commutatifs.

C'est le cas en particulier de la méthode développée en 1847 par l'astronome français Le Verrier améliorée, un siècle plus tard, par Souriau, Frame et Faddeev qui l'utilisent pour le calcul des déterminants, du polynôme caractéristique, pour l'inversion des matrices, et pour la résolution des systèmes linéaires. Cette méthode s'est avérée porteuse d'un algorithme très bien adapté au calcul parallèle, dû à Csanky, qui en 1976 a construit, dans le cas d'un anneau commutatif contenant le corps des rationnels, une famille de circuits arithmétiques, pour calculer en $\mathcal{O}(\log^2 n)$ étapes parallèles les coefficients du polynôme caractéristique.

Une autre méthode, dite *de partitionnement* ([Gas] pp. 291–298) et attribuée à Samuelson [79] (1942), a eu un regain d'intérêt avec l'algorithme de Berkowitz [6], qui fournit un calcul rapide, parallèle et sans division, du polynôme caractéristique. Cet algorithme a permis de généraliser aux anneaux commutatifs arbitraires le résultat de Csanky concernant la complexité parallèle, par une voie tout à fait différente. Nous en présenterons une version parallèle améliorée (section 10.2).

La version séquentielle la plus simple de l'algorithme de Berkowitz n'utilise pas de produits de matrices mais seulement des produits d'une matrice par un vecteur.

Elle s'est avérée tout à fait efficace sur les ordinateurs usuels, et particulièrement bien adaptée au cas des matrices creuses.

Nous présentons dans cet ouvrage les principaux algorithmes en algèbre linéaire, et donnons plus particulièrement un aperçu détaillé des différentes méthodes utilisées pour le calcul du polynôme caractéristique, avec des résultats récents.

L'intérêt porté au polynôme caractéristique d'une matrice est justifié par le fait que la détermination de ses coefficients suffit à connaître le déterminant de cette matrice et à calculer son adjointe. Dans le cas des corps cela permet de calculer son inverse et de résoudre les systèmes d'équations linéaires. Il réside également dans les renseignements que cela donne sur une forme quadratique, comme par exemple sa signature dans le cas du corps des réels.

Plan de l'ouvrage

Nous faisons quelques rappels d'algèbre linéaire dans le chapitre 1.

Le chapitre 2 contient quelques méthodes classiques couramment utilisées pour le calcul du polynôme caractéristique : l'algorithme de Jordan-Bareiss, la méthode de Hessenberg, la méthode d'interpolation de Lagrange, l'algorithme de Le Verrier et son amélioration par Souriau-Faddeev-Frame, la méthode de Samuelson modifiée à la Berkowitz, en général la plus efficace, la méthode de Chistov qui a des performances voisines, et enfin des méthodes reliées aux suites récurrentes linéaires, les plus efficaces sur les corps finis.

Le chapitre 3 développe le formalisme des circuits arithmétiques (ou programmes d'évaluation) pour une description formelle des calculs algébriques. Nous y expliquons la technique importante d'élimination des divisions, elle aussi inventée par Strassen.

Dans le chapitre 4 nous donnons un aperçu des principales notions de complexité les plus couramment utilisées. Ces notions constituent une tentative de théoriser les calculs sur ordinateur, leur temps d'exécution et l'espace mémoire qu'ils occupent.

Dans le chapitre 5 nous expliquons la stratégie générale « diviser pour gagner », bien adaptée au calcul parallèle. Nous donnons quelques exemples de base.

Le chapitre 6 est consacré à la multiplication rapide des polynômes, avec la méthode de Karatsuba et la Transformée de Fourier Discrète.

Le chapitre 7 est consacré à la multiplication rapide des matrices. Nous y abordons notamment les notions fondamentales de complexité bilinéaire, de rang tensoriel et de calculs bilinéaires approximatifs.

Le chapitre 8 est consacré à des algorithmes dans lesquels intervient la multiplication rapide des matrices, mais sans que l'ensemble de l'algorithme soit bien adapté au calcul parallèle.

On obtient ainsi en général les procédures les plus rapides connues en ce qui concerne le *temps séquentiel asymptotique*, pour la plupart des problèmes classiques liés à l'algèbre linéaire. Ces performances sont en général obtenues uniquement sur les corps. Seule la dernière section du chapitre, consacrée à l'algorithme de Kaltofen-Wiedemann concerne le calcul sur un anneau commutatif arbitraire.

Le chapitre 9 présente les parallélisations de la méthode de Le Verrier, qui s'appliquent dans tout anneau commutatif où les entiers sont non diviseurs de zéro et où la division par un entier, quand elle est possible, est explicite.

Le chapitre 10 est consacré aux méthodes parallèles de Chistov et de Berkowitz qui s'appliquent en toute généralité.

Le chapitre 11 présente tout d'abord quelques tableaux récapitulatifs des complexités des différents algorithmes étudiés, séquentiels ou parallèles, pour le calcul du déterminant et celui du polynôme caractéristique. Nous donnons ensuite les résultats des tests expérimentaux concernant quelques méthodes séquentielles du calcul du polynôme caractéristique. Ces résultats montrent des performances supérieures pour les algorithmes de Chistov et de Berkowitz avec un léger avantage pour ce dernier.

Les deux derniers chapitres sont consacrés aux travaux de Valiant sur un analogue algébrique de la conjecture $\mathcal{P} \neq \mathcal{NP}$, dans lesquels le déterminant et le permanent occupent une place centrale. Bien qu'on ait très peu d'idées sur la manière de résoudre la conjecture de Valiant $\mathcal{VP} \neq \mathcal{VNP}$, celle-ci semble quand même moins hors de portée que la conjecture algorithmique dont elle s'inspire.

L'annexe contient les codes MAPLE des algorithmes expérimentés. Nous avons choisi le logiciel de Calcul Formel MAPLE essentiellement pour des raisons de commodité. Le langage de programmation qui lui est rattaché est proche de celui de nombreux autres langages classiques, permettant de définir et de présenter de manière lisible et efficace les algorithmes considérés. Les autres langages de calcul formel généralistes auraient pu aussi bien faire l'affaire. Il n'y aura d'ailleurs aucun mal à

implémenter dans un de ces langages les algorithmes présentés dans ce livre. Une liste récapitulative en est donnée dans la table page 355.

L'esprit dans lequel est écrit cet ouvrage

Nous avons en général donné des preuves complètes de nos résultats, en accordant une grande place aux exemples. Mais il nous est aussi arrivé de ne donner qu'une idée de la preuve, ou de ne la donner complètement que sur un exemple, ou de renvoyer à une référence. Nous assumons très consciemment ce que nous avons sacrifié de la rigueur formelle au profit de la compréhension de « ce qui se passe ». Nous avons essayé de donner dessins et figures pour illustrer notre texte, tout en ayant conscience d'en avoir fait bien trop peu.

Nous avons aussi essayé de rapprocher cet exposé de la pratique concrète des algorithmes, en développant chaque fois que nous l'avons pu des calculs de complexité dans lesquels nous explicitons les constantes « cachées dans le grand $\mathcal{O}$ », sans la connaissance desquelles les résultats théoriques n'ont pas de réelle portée pratique, et peuvent être trompeurs.

Le niveau requis pour lire ce livre est seulement une bonne familiarité avec l'algèbre linéaire. Le mieux serait évidemment d'avoir lu auparavant cette perle rare qu'est le livre de Gantmacher [Gan]. On peut recommander aussi le grand classique (toujours disponible) [LT] de Lancaster & Tismenetsky. Il est naturellement préférable, mais pas indispensable, d'avoir une idée des concepts de base de la complexité binaire pour lesquels nous recommandons les ouvrages [BDG] et [Ste].

Enfin, sur les algorithmes en général, si vous n'avez pas lu le livre de Knuth [Knu] parce que vous comprenez mal l'anglais ou que vous êtes plutôt habitués à la langue de Voltaire, avant même de commencer la lecture de notre ouvrage, écrivez une lettre à tous les éditeurs scientifiques en leur demandant par quelle aberration la traduction en français n'a pas encore été faite.

Pour aller au delà en Calcul Formel nous recommandons les livres de von zur Gathen & Gerhard [GG], Bini & Pan [BP], Bürgisser, Clausen & Shokrollahi [BCS], Bürgisser [Bur] et le Handbook of Computer Algebra [GKW].

Nous espérons que notre livre contribuera à mieux faire saisir l'importance de la complexité algébrique à un moment où les mathématiques constructives et les solutions algorithmiques se développent de manière rapide et commencent à occuper de plus en plus une place essentielle

dans l'enseignement des Mathématiques, de l'Informatique et des Sciences de l'ingénieur.

Remerciements Nous remercions Marie-Françoise Roy et Gilles Villard pour leur relecture attentive et leurs suggestions pertinentes, ainsi que Peter Bürgisser pour son aide concernant les deux derniers chapitres. Et enfin François Pétiard qui nous a fait bénéficier avec une patience infinie de son expertise en LaTeX.

Table des matières