Complexité de calculs sur les matrices entières et polynomiales*

Gilles Villard

Cnrs Lip - Ens Lyon

http://www.ens-lyon.fr/ gvillard

*Journées Nationales de Calcul Formel, Luminy, 24 janvier 2003.

Problème :

Étude des **complexités** de problèmes de base en **algèbre linéaire** exacte.

- ▷ Complexité en temps *i.e.* algorithmes les plus rapides ;
- ▷ Algorithmes déterministes ou probabilistes Las Vegas.

Modèles.

Complexité algébrique.

K un corps commutatif, opérations arithmétiques $+, \times, /$

versus

Sur K$[x]$, opérations arithmétiques $+, \times, /$ dans K

Complexité binaire.

Sur $\mathbb{Z}$, nombre d'opérations sur les bits.

Motivations

$\longrightarrow$ Domaines de coefficients “concrets”

$\longrightarrow$ Complexités polynomiales et binaires mal connues.

Plan de l'exposé

1. Complexités mal connues
2. Progrès des deux dernières années
3. Matrices polynomiales
4. Réduction en colonnes
5. Procédé d'élimination des divisions de Strassen

Plan de l'exposé

1. **Complexités mal connues**
2. Progrès des deux dernières années
3. Matrices polynomiales
4. Réduction en colonnes
5. Procédé d'élimination des divisions de Strassen

Modèle algébrique sur K

Équivalence au **produit de matrice** (*straight-line*)

Produit de matrices $n \times n$ $A \times B$	**Déterminant**, rang, inversion, polynôme caractéristique, forme normale de Frobenius...
n^ω	Algorithmes RAM en $\tilde{O}(n^\omega)$

- [Strassen 69] Det $\preceq$ MM $\hookrightarrow$ MM $\preceq$ Det $\preceq$ MM
- [Strassen 73, Baur & Strassen 83] MM $\preceq$ Det

Modèle algébrique sur $\mathsf{K}[x]$ **& modèle binaire sur** $\mathbb{Z}$

Taille de la sortie : nd ou $O^{\sim}(n \log \|A\|)$

$\leftarrow$ nd points (théorème chinois) $\rightarrow$

$\uparrow$

n^{ω}

$\downarrow$

E.g. pour le déterminant

$$O^{\sim}(n^{\omega} \times nd) = O^{\sim}(n^{\omega+1}d)$$

$$O^{\sim}(n^{\omega+1} \log \|A\|)$$

Complexités guidées par la taille de la sortie, “1969–2000”

Coût binaire $\leq$ algébrique $\times$ taille en sortie

(Coût num. $\leq$ algébrique $\times$ conditionnement $\times$ erreur inverse)

Complexités guidées par la taille de la sortie

$\downarrow$

Diminution du surcoût

$\downarrow$

Déterminant (et forme de Smith) sur $\mathbb{Z}$ en $\tilde{O}(n^{2+\omega/2} \log \|A\|)$,
soit $\sqrt{n} \times n^3 \log \|A\|$ au lieu de $n \times n^3 \log \|A\|$
[Eberly, Giesbrecht, Villard 2000 (Kaltofen 1992)].

Commentaires *a posteriori* :

Gauss/Bareiss ou le théorème chinois d'emblée :
opérations sur des grandes tailles dès le départ.

Graal :

Différer les calculs sur les grandes tailles

Plus pragmatiquement :

Utiliser des algorithmes par blocs (factoriser).

Un schéma "idéal"

$\hookrightarrow n/2$ étapes d'élimination en ne multipliant la taille que par 2 ?

$$NA = T$$

$$\begin{bmatrix} \times & \times \end{bmatrix} \begin{bmatrix} \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} -85 & -55 & -37 & -35 \\ 49 & 63 & 57 & -59 \\ 43 & -62 & 77 & 66 \\ -50 & -12 & -18 & 31 \\ -91 & -47 & -61 & 41 \\ 94 & 83 & -86 & 23 \\ -53 & 85 & 49 & 78 \\ -86 & 30 & 80 & 72 \end{bmatrix}$$

Élimination de Gauss (complément de Schur) :

$$N_g = \begin{bmatrix} 7646610 & -17525750 & -3967680 & 29755220 & \ldots\ldots \\ -15181842 & 13894262 & 0 & -40184660 & \ldots\ldots \\ -2804568 & 4081928 & 0 & 18871120 & \ldots\ldots \\ 4368828 & -4023028 & 0 & 35835160 & \ldots\ldots \end{bmatrix}$$

Élimination de Gauss (complément de Schur) :

$$N_g = \begin{bmatrix} 7646610 & -17525750 & -3967680 & 29755220 & \dots\dots \\ -15181842 & 13894262 & 0 & -40184660 & \dots\dots \\ -2804568 & 4081928 & 0 & 18871120 & \dots\dots \\ 4368828 & -4023028 & 0 & 35835160 & \dots\dots \end{bmatrix}$$

alors que l'on peut construire la base

$$N = \begin{bmatrix} -25 & -32 & -16 & -38 & 1 & -30 & 32 & -33 \\ -27 & -68 & -43 & 23 & -71 & -1 & -55 & 61 \\ 106 & -43 & 28 & -95 & -50 & 30 & 53 & -7 \\ -23 & -25 & -12 & 182 & -90 & -40 & 36 & -74 \end{bmatrix}$$

$$n^2 \sum_{i=1}^{n} i \approx n^4$$

$$\sum_{i=1}^{\log n} (\frac{n}{2^i})^3 \, 2^i \approx n^3$$

Un schéma idéal :

- En partie atteint dans le cas polynomial,
  [Storjohann 2002] [Jeannerod & Villard 2002].

- Complexité "indépendante" de la "taille individuelle" en sortie ?

- (Complexité "indépendante" du conditionnement ?)

Plan de l'exposé

1. Complexités mal connues
2. **Progrès des deux dernières années**
3. Matrices polynomiales
4. Réduction en colonnes
5. Procédé d'élimination des divisions de Strassen

Suppression du facteur n / Modèle binaire sur $\mathbb{Z}$

	Facteur	Total
Déterminant (Las Vegas)	$n^{1/5}$	$n^{2.7} \log \|\|A\|\|$
Forme de Smith (Las Vegas)	$n^{1/5}$	$n^{2.7} \log \|\|A\|\|$
Forme de Frobenius (Monte Carlo)	$n^{1/5}$	$n^{2.7} \log \|\|A\|\|$

(et sans division sur K pour le déterminant)

[Kaltofen & Villard, 2001, 2002].

Suppression du facteur n / Modèle algébrique sur $\mathsf{K}[x]$

	Facteur	Total
Déterminant & Smith[1] (RAM)	$\log^{\alpha} n$	$O^{\sim}(n^{\omega} d)$
Inversion & det.[2] (SLP)	$\log^{\beta} n$	$O^{\sim}(n^{3} d)$ & $O^{\sim}(n^{\omega} d)$

[1] [Storjohann 2002],

[2] [Jeannerod & Villard 2002].

Jusque-là, calculs en $\tilde{O}(n^{\omega+1}d)$ ou $\tilde{O}(n^{\omega+1}\log\|A\|)$,
algorithmes déterministes ou Las Vegas,

Déterminant, rang certifié, inversion

polynôme caractéristique, forme normale de Frobenius,

Formes normales d'Hermite, Smith et transformations.

$\longrightarrow$ **"Ré-unification" des complexités ?**

(Gauss *versus* Krylov)

Plan de l'exposé

1. Complexités mal connues
2. Progrès des deux dernières années
3. **Matrices polynomiales**
4. Réduction en colonnes
5. Procédé d'élimination des divisions de Strassen

Nota. Études sur K$[x]$, préliminaires à celles sur $\mathbb{Z}$.

Matrices polynomiales $n \times n$ de degré d

Nouvelles complexités en $n^\omega d$

Nota. Études sur $\mathsf{K}[x]$, préliminaires à celles sur $\mathbb{Z}$.

Matrices polynomiales $n \times n$ de degré d

Nouvelles complexités en $n^\omega d$

$\Downarrow$

Lien avec le produit de deux matrices de degré d ?

Définition : $\mathsf{MM}(n, d)$ et $\mathsf{Det}(n, d)$ coûts du produit et du déterminant $n \times n$ de degré d.

$$\mathsf{MM}(n,d) \preceq \mathrm{Det}(n,d)\,?$$

Le déterminant implique-t'il la multiplication ?

Sur K, **[Baur & Strassen 83]**,

$$\det A = a_{1,1} \times (\text{ mineur } (n-1) \times (n-1)) + \ldots$$

donc, les coefficients de A^*, l'adjointe,

$$a^*_{j,i} = \frac{\partial \det A}{\partial\ a_{i,j}}$$

Théorème. $\mathrm{MM}(n) \preceq\ \mathrm{Det}(n)$ (arbres de calculs ou $\textsc{Slp}$).

Preuve. Différentiation automatique en mode inverse et utilisation de $\mathrm{MM}(n) \preceq\ \mathrm{Inv}(n)$ (cf. plus loin).

Théorème. S'il existe un programme de longueur $Det(n,d)$ calculant le coefficient de degré d du déterminant de $A(x)$, alors il existe un programme de longueur $O\tilde{}(\mathrm{Det}(n,d))$ pour le produit $B(x) \times C(x)$ de degré d.

Preuve.

$$\det A = (\ldots + a_{1,1,k}\, x^k + \ldots) \times (\text{ mineur } (n-1) \times (n-1)) + \ldots$$

donc,

$$\begin{aligned} \partial \det A / \partial\, a_{i,j,k} &= x^k\, a^*_{j,i} \\ &= x^k\, a^*_{j,i,0} + x^{k+1}\, a^*_{j,i,1} + x^{k+2}\, a^*_{j,i,2} + \ldots \end{aligned}$$

or,

$$\frac{\partial \det A}{\partial\, a_{i,j,k}} = \ldots + \frac{\partial \Delta_d}{\partial\, a_{i,j,k}} x^d + \ldots$$

donc,

$$\frac{\partial \det \Delta_d}{\partial\, a_{i,j,k}} = a^*_{i,j,d-k}$$

Différentiation du terme de degré d du déterminant par rapport aux $n^2 d$ entrées de la matrice

$$\Downarrow$$

d premiers termes de l'adjointe *i.e.* $A^* \bmod x^d$.

Et,

$$A = \begin{bmatrix} I & B & \\ & I & C \\ & & I \end{bmatrix} \Rightarrow A^*(x) \equiv \begin{bmatrix} I & -B & B(x)C(x) \\ & I & -C \\ & & I \end{bmatrix} \bmod x^d$$

$+$ Aménagement technique pour $B(x)C(x) \bmod x^{2d+1}$.

Donc, $\mathsf{MM}(n,d) \preceq \mathrm{Det}(n,d)$.

Inversement, le déterminant se calcule en $O^{\sim}(\mathsf{MM}'(n,d))$ où $\mathsf{MM}'(n,d)$ est reliée à $\mathsf{MM}(n/2^i, 2^i d)$ ($\log n$ étapes)
[Storjohann 2002, Jeannerod & Villard 2002].

$$\mathsf{MM}(n,d) \preceq \mathrm{Det}(n,d) \preceq O^{\sim}(\mathsf{MM}'(n,d)).$$

Plan de l'exposé

1. Complexités mal connues
2. Progrès des deux dernières années
3. Matrices polynomiales
4. **Réduction en colonnes**
5. Procédé d'élimination des divisions de Strassen

Réduction en colonnes - définition.

$$A(x) = \begin{bmatrix} x+1 & x^2 \\ x^2 & x^3+x^2+1 \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} x+1 & x \\ x^2 & x^2+1 \end{bmatrix} \rightarrow \quad C(x) = \begin{bmatrix} x+1 & 1 \\ x^2 & 1 \end{bmatrix}$$

Réduction en colonnes - définition.

$$A(x) = \begin{bmatrix} x+1 & x^2 \\ x^2 & x^3+x^2+1 \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} x+1 & x \\ x^2 & x^2+1 \end{bmatrix} \rightarrow \quad C(x) = \begin{bmatrix} x+1 & 1 \\ x^2 & 1 \end{bmatrix}$$

Définition. La matrice de tête de $C = AU$ en colonne $[C]_l$ est de rang maximal.

Conséquence. Les colonnes de C forment une base de degrés minimaux du K$[x]$-module correspondant, dite **base minimale**.

Difficulté.

Degrés inférieurs à d dans la matrice mais grands degrés dans la transformation.

Exemple.

$$\begin{bmatrix} 1 & x & 0 & 0 \\ 0 & 1 & x & 0 \\ 0 & 0 & 1 & x \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -x & x^2 & -x^3 \\ 0 & 1 & -x & x^2 \\ 0 & 0 & 1 & -x \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Autre caractérisation sur $A^{-1}(x)$.

Description de fractions matricielles $P(x)Q^{-1}(x)$ *i.e.* "$Q(x)/P(x)$"

La fraction $A^{-1}(x)$ n'est pas strictement propre,

$$\lim_{x\to\infty} A^{-1}(x) \neq 0.$$

Conséquence, ex :

$$\frac{x^6 + 2\,x^5 - x^3 - 1}{x^2 + x + 1}$$

Padé $\dfrac{[2]}{[2]} = \dfrac{-1 - x^2}{1 + x + 2\,x^2}$, $\dfrac{[4]}{[2]} = \dfrac{\dots}{1 + \frac{1}{4}\,x + \frac{5}{4}\,x^2}$, $\dfrac{[6]}{[2]} = \dfrac{\dots}{x^2 + x + 1}$

Nouvel algorithme.

Réduction en colonnes

Entrée : $A \in \mathsf{K}[x]^{n \times n}$ de degré d

Sortie : $C = AU$ réduite en colonnes

> Calculer $2d$ termes du dév. de A^{-1} après l'ordre $(n-1)d$
> Reconstruire un approximant de Padé matriciel BC^{-1}
> Retourner C

I) $2d$ termes d'ordres supérieurs à $(n-1)d$ de $A^{-1}(x)$

["**High-order lifting**", Storjohann 2002]

Avec $A^{-1}(x) = A_0 + xA_1 + \ldots + x^i A^i + \ldots$,

$\log n$ étapes, calcul de "tranches" du développement :

$$
\left|
\begin{array}{l}
A_0, \ldots, A_d \\
\qquad A_d, \ldots, A_{2d} \\
\qquad\qquad A_{3d}, \ldots, A_{4d} \\
\qquad\qquad\qquad \ldots \\
\qquad\qquad\qquad\qquad A_{(n-1)d}, \ldots, A_{nd}
\end{array}
\right.
$$

$\hookrightarrow$ Essentiellement en $O(\mathsf{MM}(n,d)\log n)$.

II) Reconstruction de la fraction avec matrice dénominateur réduite

[Giorgi, Jeannerod & Villard 2002]

$\hookrightarrow$ Essentiellement en $\mathsf{MM}'(n,d) \approx O(\mathsf{MM}(n,d)\log d)$.

Algorithme de type **Knuth / Schönhage pour polynômes matriciels**

[Beckermann & Labahn 94] [Coppersmith 94]

[Thomé 2001] [Kaltofen & Villard 2001]

$$\frac{B_0}{C_0} \equiv H \bmod x \quad \rightarrow \quad \frac{B_h}{C_h} \equiv H \bmod x^h \quad \rightarrow \quad \frac{B_{2h}}{C_{2h}} \equiv H \bmod x^{2h}$$

$$M_{2h} = \overline{M}_{2h} \times M_h.$$

Théorème. Une matrice inversible peut être colonne réduite par un algorithme Las Vegas en $O(\mathsf{MM}'(n,d)\log n)$ ou $O^{\sim}(n^{\omega}d)$ opérations dans K.

Plan de l'exposé

1. Complexités mal connues
2. Progrès des deux dernières années
3. Matrices polynomiales
4. Réduction en colonnes
5. **Procédé d'élimination des divisions de Strassen**

Modèle algébrique sur un anneau (sans division)

Procédé d'élimination des divisions de Strassen,

$$1/(1-za) \equiv 1 + az + a^2z^2 + \ldots a^n z^n \bmod z^{n+1}.$$

I) Départ : un programme $\mathcal{P}$ sur un corps K avec divisions

II) Homotopie en utilisant un point M où $\mathcal{P}$ ne divise que par 1

III) Exécution de $\mathcal{P}$ en plongeant K dans K$[[z]]$.

Modèle algébrique sur un anneau (sans division)

Procédé d'élimination des divisions de Strassen,

$$1/(1-za) \equiv 1 + az + a^2z^2 + \ldots a^n z^n \bmod z^{n+1}.$$

I) Départ : un algorithme $\mathcal{P}$ sur un corps K avec divisions

II) Homotopie en utilisant un point M où $\mathcal{P}$ ne divise que par 1

III) Exécution de $\mathcal{P}$ en plongeant K dans K$[[z]]$.

Illustration :

$$\det A = \left(\det\left((1-z)M + zA\right) \bmod z^{n+1}\right)_{|z=1}$$

Procédé de Strassen

Exécution d'un algorithme sur K, et sur K$[[x]]$

$\downarrow$

Analogie fortuite

$\downarrow$

Limiter le surcoût lié à la taille séries formelles

pour ne pas avoir $n \times n^{\omega}$

Différence importante : ici, **un algorithme sous-jacent sur** K

Conclusion

$\hookrightarrow$ Gagner un facteur n n'est sans doute pas anecdotique,

cela reflète une compréhension incomplète des problèmes

dans leur ensemble.

$\hookrightarrow$ Que peut-on calculer en temps $\tilde{O}(n^\omega d)$ ou $\tilde{O}(n^\omega \log \|A\|)$?