ALA, 31 mai 2001 – Rabat.

Evaluation du Signe et Calcul du Déterminant d'une Matrice à Coefficients Entiers

Gilles Villard

Cnrs - Lip ENS Lyon

$$A = \begin{bmatrix} 197995567 & 62556 & 975255 & 97 \\ 99292191 & 953447 & 875753 & 87 \\ 29 & 98565213 & 98763653 & 9924 \\ 25 & 9932 & 9952 & 1 \end{bmatrix} \in \mathbb{Z}^{4\times 4}$$

- Calcul exact :
  $\rightarrow \quad \det(A) = 30$

- Simple précision flottante (QR) :
  $\rightarrow \quad \det(A) = 3.8110^{14}$

- Double précision flottante (QR) :
  $\rightarrow \quad \det(A) = -233.0$

- 60 bits (QR) :
  $\rightarrow \quad \det(A) = 11.0$

1. Quel est le coût du calcul du signe ?
(calculs approchés éventuellement)

2. Quel est le coût du calcul du déterminant ?

En conséquence :
Sur combien de bits calculer au moins ?

- Critère considéré (pour cet exposé) :

$\boxed{\textbf{Complexité en temps}}$

Paramètres en entrée

$$A = \begin{bmatrix} 197995567 & 62556 & 975255 & 97 \\ 99292191 & 953447 & 875753 & 87 \\ 29 & 98565213 & 98763653 & 9924 \\ 25 & 9932 & 9952 & 1 \end{bmatrix}$$

- **La dimension** : n , $n = 4$.
- **Le nombre de bits en entrée** : β , $\beta = 28$,
  $\| A \|_\infty \leq 2^\beta = 2^{28}$.

$$\boxed{\mathcal{S}_{\mathrm{det}}(n, \beta)? \;\; \mathcal{D}(n, \beta)?}$$

- **Paramètres implicites** :
  - nombre de conditionnement,
  - défaut d'orthogonalité,
  - invariants de structure,

  ...

Motivations

- Prédicats en géométrie.
- nombres algébriques réels.
- Calcul symbolique.

Plan de la présentation

I/ Aspects algébriques

II/ Coûts «classiques»

III/ Signe : une approche numérique

IV/ Systèmes linéaires : une approche exacte

V/ «Sans division» : une autre approche exacte

VI/ Synthèse.

Rappel : on travaille exactement avec tous les nombres d'un corps donné (ex $\mathbb{Q}$) sans considération de taille, de nombre de bits ou de précision.

Complexité des fractions rationnelles,

- ▷ MM : produit de matrices,
- ▷ Det : déterminant.
- ▷ SysLin : résolution d'un système inversible.

- [Strassen] :

$$\mathbf{MM} \Rightarrow \mathbf{Det}, \quad \mathrm{MM} \Rightarrow \mathrm{SysLin}$$

Preuve : factorisation récursive.

- [Strassen, Baur & Strassen] :

$$\mathbf{Det} \Rightarrow \mathbf{Matrice\ adjointe}$$

Preuve :

$$\frac{\partial}{\partial a_{i,j}} \det A = \frac{(A^{-1})_{i,j}}{\det A} = a^*_{i,j}$$

+ calcul de toutes les dérivées partielles.

$$\text{Matrice adjointe} \Rightarrow \text{MM}$$

Preuve :

$$\begin{bmatrix} 1 & A & 0 \\ 0 & 1 & B \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -A & AB \\ 0 & 1 & -B \\ 0 & 0 & 1 \end{bmatrix}.$$

Donc :

$$\mathbf{SysLin} \Leftarrow \mathbf{Det} \Longleftrightarrow \mathbf{MM}$$

Modèle de complexité binaire

- Determinant d'une matrice $n \times n$ sur $\mathbb{Z}$

$$A = \begin{bmatrix} 43 & -62 & 77 & 66 & 54 & -5 & 99 \\ -61 & -50 & -12 & -18 & 31 & -26 & -62 \\ 1 & -47 & -91 & -47 & -61 & 41 & -58 \\ -90 & 53 & -1 & 94 & 83 & -86 & 23 \\ -84 & 19 & -50 & 88 & -53 & 85 & 49 \\ 78 & 17 & 72 & -99 & -85 & -86 & 30 \\ 80 & 72 & 66 & -29 & -91 & -53 & -19 \end{bmatrix}$$

$$\det \quad A = -222570833995966.$$

$$\det A \leq (n/2)\log n + n\beta$$

soit

$$\boxed{\text{Taille} \det A = \log|\det A| \approx \tilde{O}(n\beta) \text{ (Hadamard)}}$$

NB. Dans le cas binaire, il n'existe pas vraiment d'analogue aux réductions entre les différents problèmes comme vu en complexité algébrique.

Complexité binaire «classique» :

Coût binaire = taille du déterminant × coût arithmétique.

$$\boxed{\rightarrow \quad \text{Coût en } O\tilde{}(n^3 \times n\beta) = O\tilde{}(n^4\beta)}$$

Algorithmes correspondants :

- élimination de Gauss + grands entiers,
- arithmétique multi-modulaire,

...

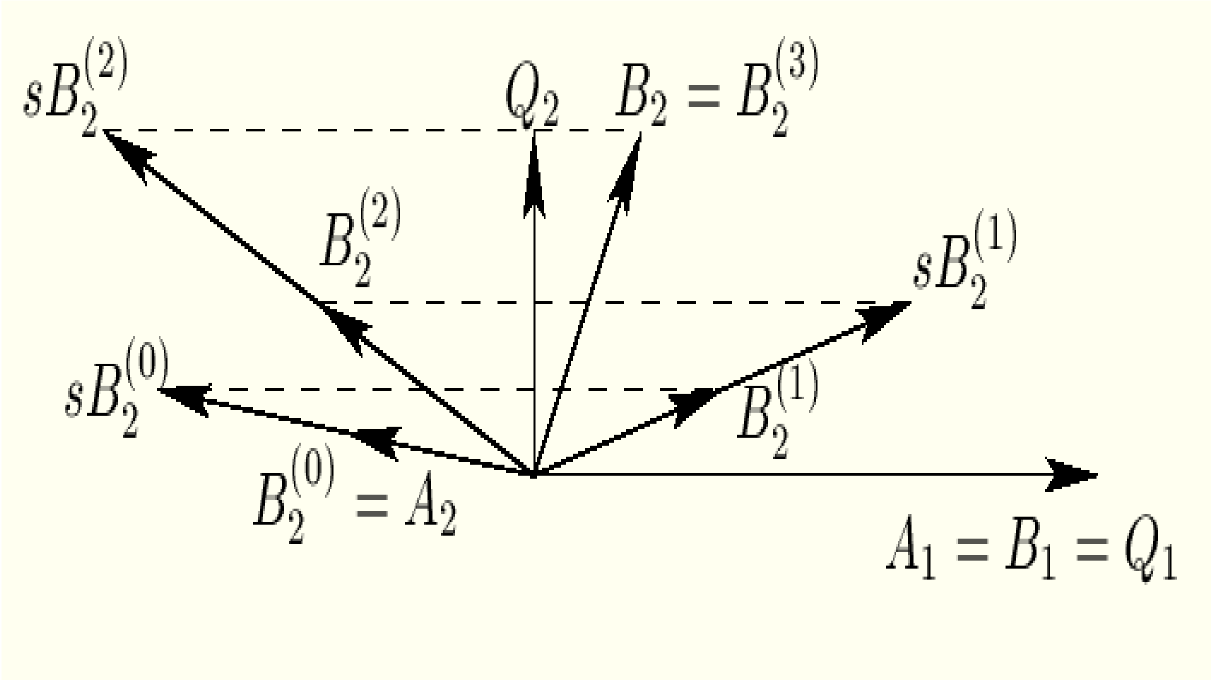
L'algorithme de Clarkson et ses prolongements

[Clarkson] [Brönnimann & Yvinec]

1. Calcul d'une matrice B «plus orthogonale» que A
   et de même signe de déterminant
2. Calcul du signe du déterminant de B
   par exemple par décomposition LU

Clarkson(A,n)
  pour i de 1 à n
    $B_i := A_i$
    tant que B_i «pas presque orthogonal» à $Q_1, \ldots, Q_{i-1}$
      $B_i :=$ Réduction + changement d'échelle(B_i)
    fin tant que
    $Q_i :=$ Gram-Schmidt(B_i)
  fin pour
  $det :=$ LU(Q)
  retourner signe(det)

Réduction + changement d'échelle
en trois étapes



Précision suffisante : $\beta + O(n)$ bits.

Coût :

$$\mathcal{S}_{\det}(n, \beta) \leq \tilde{O}((n^3 + n^2\delta(A)) \cdot (n + \beta))$$

Le **coût adaptatif** fait apparaître un **défaut d'orthogonalité**. Le nombre d'itérations de la première phase dépend de :

$$\delta(A) = \log \frac{1}{\Delta(A)} = \log \frac{\prod_{i=1}^{n} \| a_i \|}{\mathrm{vol}(a_1, \ldots a_n)}$$

$\delta(A) = 1$ si les colonnes de A sont orthogonales.
$\delta(A) = O(n\beta)$ dans le pire cas (matrice unimodulaire).

Pire cas équivalent à l'approche classique.

- **Cas favorable** à Clarkson

$$\mathcal{S}_{\mathrm{det}}(n, \beta) \leq O\tilde{}((n^3 + n^2\delta(A)) \cdot (n + \beta))$$

$$\downarrow$$

Avec $\delta(A) = O(n)$ (A est «bien conditionnée»)

$$\mathcal{S}_{\mathrm{det}}(n, \beta) \leq O\tilde{}(n^4 + n^3\beta).$$

Tirer parti du fait que résoudre un système est *a priori* «plus simple» [Pan]

[Abbott,Bronstein & Mulders][Eberly,Giesbrecht & Villard]

$Ax = b$, par les **règles de Cramer** :

$$x_j = \frac{|A, b|_j}{\det A}$$

$$\text{où} \quad |A, b|_j = \sum_i (-1)^{i+j} b_i \, A \begin{pmatrix} \{1 \ldots n\} \setminus \{i\} \\ \{1 \ldots n\} \setminus \{j\} \end{pmatrix}.$$

Remarque importante :

Suivant la structure de A (forme normale de Smith), il se peut que quelquesoit le vecteur b second membre, la fraction de Cramer se simplifie toujours en :

$$x_j = \frac{|A, b|'_j}{\sigma_A}$$

Exemple :

A l'aide d'un système

$$Ax = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 12 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$$

on ne peut avoir que

$$x = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4/4 \\ b_5/12 \end{bmatrix}$$

et donc au mieux le diviseur $\sigma_A = 12 \neq \det A = 24$.

Deux types d'algorithmes.

Algorithme 1.
[Abbott,Bronstein & Mulders]

1. Calculer un diviseur σ_A du déterminant
   par résolution de $Ax = b$ (b aléatoire)

2. Calculer le facteur manquant $\det(A)/\sigma_A$
   par la méthode classique

Pire cas équivalent à l'approche classique.

- **Cas favorable**

$$\log \frac{\prod_{i=1}^{n} \parallel a_i \parallel}{|\sigma|} = O(n)$$

$$\downarrow$$

$$\mathcal{D}(n,\beta) \leq O\tilde{}(n^4 + n^3\beta^2).$$

Algorithme 2 (probabiliste).
[Eberly, Giesbrecht & Villard]

1. Calculer un diviseur σ_A du déterminant
   par résolution de $Ax = b$ (b aléatoire)

2. Calculer le facteur manquant $\det(A)/\sigma_A$
   en ré-appliquant la méthode à des
   perturbations successives de la matrice.

Le nombre de perturbations à effectuer dépend d'un paramètre ϕ de structure de A (forme normale de Smith) :

$$1 \leq \phi \leq O(\sqrt{n\beta}).$$

Résultat :

Même dans le pire cas, la formule

Coût binaire $=$ taille du déterminant $\times$ coût arithmétique.

est remplacée par

$$\boxed{\text{Coût binaire } = \sqrt{n\beta} \times \text{ coût arithmétique.}}$$

- Etude du **cas moyen** (et cas favorable) :

$$\mathsf{E}(\phi) = \frac{\log n}{\beta}$$

donc le coût en moyenne :

$$\mathsf{E}(\mathcal{D}(n,\beta)) \leq \tilde{O}(\mathsf{E}(\phi) \times (n^3\beta^2)) = \tilde{O}(n^3\beta).$$

V/ «Sans division» : une autre approche exacte

[Wiedemann] [Kaltofen & Villard]

1. Approche Lanczos-Wiedemann «classique»

$$I, A, A^2, \ldots, A^{d-1}, \quad \text{le premier vecteur dépendant } A^d$$

$$u^t v, u^t A v, u^t A^2 v, \ldots, u^t A^{2d} v \quad \text{Berlekamp-Massey}$$

$$\downarrow$$

Polynôme minimal de A

$$\pi(x) = x^d + p_{d-1}x^{d-1} + \ldots + p_1 x + p_0$$

$$\downarrow$$

B «conditionnée», $B = DA$ a des valeurs propres distinctes

Polynôme caractéristique = polynôme minimal

$$\chi(x) = x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0$$

$$\downarrow$$

$$\boxed{\det B = c_0.}$$

Problème :

$$\text{Calcul de } v^t u, v^t Au, v^t A^2 u, \ldots, v^t A^n u\,?$$

En tenant compte de la taille.

Une approche naïve conduit de nouveau à

$$\text{Coût} = \text{taille} \times \text{arithmétique}$$

2. A pas de bébé et pas de géant

[Shanks] Logarithmes discrets

[Kaltofen] Déterminant

[Kaltofen & Shoup] Factorisation de polynômes

$$u, Au, A^2u, \ldots, A^n u$$

$$\left.\begin{array}{l} n \text{ produits par } A, \text{ chacun en } O(n^2) \\ \text{taille } +1 \text{ à chaque produit} \\ \text{donc taille en } O(n) \end{array}\right\} \text{ coût en } O(n^4)$$

Principe :

supprimer les deux boucles imbriquées
pour éviter le produit taille $\times$ nb ops.

I/ Le bébé : $u, Au, A^2u, \ldots, A^{\sqrt{n}}u$

II/ Le géant : $G = A^{\sqrt{n}}$

III/ Le bébé : $v, v^tG, v^tG^2, v^tG^{\sqrt{n}}$

IV/ Conclusion : former tous les produits $(v^tG^i)\cdot A^ju$ pour les

$$v^tA^ku, k = 1, \ldots, n.$$

Coût : nb produits × coût d'un produit × taille

I/ $\sqrt{n} \times n^2 \times \sqrt{n}$

II/ $\log n \times n^3 \times \sqrt{n}$

III/ $\sqrt{n} \times n^2 \times n$

IV/ $n \times n \times n$

$$\boxed{\text{Total : } \tilde{O}(n^{3.5}).}$$

Gain d'un facteur $\sqrt{n}$ supplémentaire

3. Block Lanczos-Coppersmith / Polynôme minimal matriciel

$$U, V \in \mathbb{Z}^{n\times m},\;\; U^tV,\; U^tAV,\; U^tA^2V, \ldots,\; U^tA^{2d}V$$

$$(U^tA^dV) + (U^tA^{d-1}V)C_{d-1} + \ldots + (U^tAV)C_1 + (U^tV) = 0$$

$$\downarrow$$

$$x^dI + x^{d-1}C_{d-1} + \ldots + xC_1 + C_0 \quad \in \mathbb{Q}^{m\times m}$$

Polynôme matriciel minimal :

Calculé par pgcd de matrice
 ou par approximants de Padé matriciels
 ou par résolution d'un système Hankel par blocs
 ...

Polynôme matriciel minimal

Autre format pour la «même» information sur le spectre

$$A - xI = \begin{bmatrix} 1/3 & -2 & -8/3 & -8/3 \\ 1/3 & 2 & 4/3 & 4/3 \\ -1 & 0 & 2 & -1 \\ 8/3 & 2 & 8/3 & \frac{17}{3} \end{bmatrix} - xI$$

$$\Updownarrow$$

$$C(x) = \begin{bmatrix} -3\,x - 8/5 + x^2 & 6\,x - 24 \\ -3/5\,x + 9/5 & -7\,x + 12 + x^2 \end{bmatrix}$$

$$\Updownarrow$$

$$\chi(x) = x^4 - 10\,x^3 + 35\,x^2 - 50\,x + 24$$

Exemple :

$$A = \begin{bmatrix} 1/3 & -2 & -8/3 & -8/3 \\ 1/3 & 2 & 4/3 & 4/3 \\ -1 & 0 & 2 & -1 \\ 8/3 & 2 & 8/3 & \frac{17}{3} \end{bmatrix}$$

Polynôme minimal (scalaire)

$u^t A^i v$: $\quad 2, 16/3, 18, \frac{202}{3}, 262, \frac{3106}{3}, 4118, \frac{49282}{3}, 65622\ldots$

$+$ Berlekamp Massey $\rightarrow \chi(x) = x^4 - 10\,x^3 + 35\,x^2 - 50\,x + 24.$

Polynôme minimal matriciel

$U^t A^i V$: $\quad \begin{bmatrix} 2 & 4 \\ 4 & 16 \end{bmatrix}, \begin{bmatrix} 16/3 & 10 \\ \frac{32}{3} & 44 \end{bmatrix}, \begin{bmatrix} 18 & 38 \\ 36 & 148 \end{bmatrix}, \begin{bmatrix} \frac{202}{3} & 166 \\ \frac{404}{3} & 548 \end{bmatrix}, \begin{bmatrix} 262 & 734 \\ 524 & 2116 \end{bmatrix}$

$$\begin{bmatrix} 2 & 4 \\ 4 & 16 \end{bmatrix} \cdot \begin{bmatrix} \frac{8}{5} & 24 \\ -\frac{9}{5} & -12 \end{bmatrix} + x \begin{bmatrix} \frac{16}{3} & 10 \\ \frac{32}{3} & 44 \end{bmatrix} \cdot \begin{bmatrix} 3 & -6 \\ \frac{3}{5} & 7 \end{bmatrix} - x^2 \begin{bmatrix} 18 & 38 \\ 36 & 148 \end{bmatrix} \cdot \begin{bmatrix} 1 & \\ 0 & 1 \end{bmatrix} = 0$$

$+$ Matrix Padé $\rightarrow$ $C(x) = \begin{bmatrix} -3\,x - 8/5 + x^2 & 6\,x - 24 \\ -3/5\,x + 9/5 & -7\,x + 12 + x^2 \end{bmatrix}$

- **Propriété :** $\det C(x) = \chi(x) \;\Rightarrow\; \det A = \det C(0).$

Exemple : Calcul du déterminant – [Kaltofen & Villard].

`Entrée :`

$$A = \begin{bmatrix} 2 & -1 & -5 & -3 & -4 & 2 \\ -10 & -7 & 7 & 7 & 2 & -9 \\ -4 & -3 & 11 & 7 & 5 & -5 \\ -3 & -4 & -6 & -2 & -5 & -2 \\ 12 & 11 & -11 & -10 & -2 & 12 \\ 5 & 5 & 0 & -2 & 2 & 4 \end{bmatrix}.$$

`Sortie :` $\det A$?

```
Etape 1.      /* Conditionnement */
```

Le polynôme minimal de A est différent de son polynôme caractéristique, on peut le voir sur une forme semblable :

$$F_A = \begin{bmatrix} 0 & 1 & & & & \\ 1 & 2 & & & & \\ & & 0 & 1 & & \\ & & 1 & 2 & & \\ & & & & 0 & 1 \\ & & & & 1 & 2 \end{bmatrix}$$

$$\downarrow$$

Changement d'échelle aléatoire
diag$(1, 2, 3, 4, 5, 6)$

$$\downarrow$$

La nouvelle matrice B a des valeurs propres distinctes :

$$B = \begin{bmatrix} 2 & -1 & -5 & -3 & -4 & 2 \\ -20 & -14 & 14 & 14 & 4 & -18 \\ -12 & -9 & 33 & 21 & 15 & -15 \\ -12 & -16 & -24 & -8 & -20 & -8 \\ 60 & 55 & -55 & -50 & -10 & 60 \\ 30 & 30 & 0 & -12 & 12 & 24 \end{bmatrix}, \quad F_B = \begin{bmatrix} & & & & & 720 \\ 1 & & & & & 2136 \\ & 1 & & & & -3480 \\ & & 1 & & & -1588 \\ & & & 1 & & 43 \\ & & & & 1 & 27 \end{bmatrix}.$$

```
Etape 2.
/* Calcul du polynôme minimal matriciel */
```

```
/* Génération de la suite de matrices */
```

$$C(x) = \begin{bmatrix} 13 & 8 \\ 8 & 14 \end{bmatrix}, \begin{bmatrix} 217 & 185 \\ 107 & 97 \end{bmatrix}, \begin{bmatrix} 11280 & 9937 \\ -109 & -137 \end{bmatrix}, \begin{bmatrix} 376483 & 329832 \\ 44410 & 39843 \end{bmatrix}$$

```
/* Pgcd de matrices */
```

$$\begin{bmatrix} -\frac{174528786}{3457457} - \frac{8507719079}{17287285}\,x - \frac{3160837177}{17287285}\,x^2 + x^3 & -\frac{175402998}{3457457} - \frac{7511942372}{17287285}\,x - \frac{2766671371}{17287285}\,x^2 \\ -\frac{209510760}{3457457} + \frac{2483844884}{17287285}\,x + \frac{3074596254}{17287285}\,x^2 & -\frac{161245200}{3457457} + \frac{2070367712}{17287285}\,x + \frac{2694080482}{17287285}\,x^2 + x^3 \end{bmatrix}$$

```
Etape 3.
/* Où l'on retrouve le déterminant */
```

$$\det C(0) = -720$$

On revient à A *i.e.* on «défait» le conditionnement :

$$\rightarrow \qquad \det A = -720/720 = -1.$$

- $\triangleright$ Méthode de Krylov.
- $\triangleright$ polynôme minimal matriciel.
- $\triangleright$ Elimination des divisions.
- $\triangleright$ Pas de bébé et pas de géant.
- $\triangleright$ Produit rapide de polynômes.
- $\triangleright$ Produit rapide de matrices.

$$\downarrow$$

$$\mathcal{D}(n,\beta) \leq O\tilde{}\,(\sqrt[3]{n}\cdot(n^3\beta)) \leq (n^{3.34}\beta),$$

Avec produit rapide de matrices ce coût devient :

$$\mathcal{D}(n,\beta) \leq O(n^{2.7}\beta).$$

Comparaison des complexités, $\tilde{O}$ $- A \in \mathbb{Z}^{n\times n}, \quad \log \| A \|_\infty \leq \beta$

Méthodes	Pire cas	Cas favorables
Classique	$n^4\beta + n^3\beta^2$	$n^4\beta + n^3\beta^2$
Clarkson	$(n^3 + n^2 \log \frac{1}{\Delta(A)}) \cdot (n+\beta)$ $n^4\beta + n^3\beta^2$	$n^4 + n^3\beta$
Syst. lin.	$n^3\beta^2 + \left(\log \frac{\prod_{i=1}^n \|a_i\|}{\|\sigma\|}\right) \cdot \left(n^2\beta + n^3\right)$ $n^4\beta + n^3\beta^2$	$n^4 + n^3\beta^2$
Smith	$\sqrt{\|\det(A)\|} \cdot n^3\beta^2$ $n^{3.5}\beta^{2.5}$	$\mathsf{E}(\mathcal{D}(n,\beta)) \leq n^3\beta^2$
Sans division	$n^{3.34}\beta$	$n^{3.34}\beta$

□ Pire cas :

$$n^4\beta$$

□ Cas favorable :

$$n^4 + n^3\beta$$

□ Coût moyen :

$$n^3\beta$$

□ Meilleur coût asymptotique :

$$n^{3.34}\beta \quad (n^{2.7}\beta)$$

Peut-on espérer

$$n^3\beta \quad (n^\omega\beta) \ \ ?$$