Matrices creuses et algorithme de Wiedemann

Vincent Pilaud

2007

1 Introduction

Soit $\varphi : \mathbb{N} \to \mathbb{N}$ telle que $\varphi(n) = o(n)$ lorsque $n \to \infty$. Soit $\mathbb{K}$ un corps et $m \in \mathbb{N}$. On dit qu'une matrice $A = [a_{ij}]_{(i,j)\in\{1,\dots,m\}^2}$ de $M_m(\mathbb{K})$ est *φ-creuse* si pour tout $i \in \{1, \dots, m\}$,

$$\operatorname{card}\{j \in \{1, \dots, m\} \mid a_{ij} \neq 0\} \leq \varphi(m).$$

Si A est une matrice φ-creuse de $M_m(\mathbb{K})$ et $B \in M_m(\mathbb{K})$, le calcul de AB se fait en $m^2\varphi(m)$ multiplications au lieu de m^3, de sorte que l'on va toujours privilégier la multiplication par A sur d'autres opérations matricielles possibles. On présente par exemple dans ce texte l'algorithme de Wiedemann (§ 3) qui inverse une matrice creuse A en n'utilisant que des multiplications par A. Pour cela, on montre comment l'algorithme d'Euclide étendu permet de déterminer le polynôme minimal d'une suite récurrente linéaire scalaire de degré d dont on connaît les $2d$ premiers termes (§ 2).

2 Suites récurrentes linéaires

2.1 Définition

Soit $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel. On considère l'espace vectoriel $\mathcal{S} = E^{\mathbb{N}}$ des suites à valeurs dans E. On note $\delta : \mathcal{S} \to \mathcal{S}$ l'*opérateur de décalage* défini par $\delta(u)_n = u_{n+1}$, pour toute suite $u = (u_n)_{n\in\mathbb{N}} \in \mathcal{S}$. Cet opérateur est clairement linéaire. On dit qu'une suite $u = (u_n)_{n\in\mathbb{N}} \in \mathcal{S}$ est une *suite récurrente linéaire* s'il existe $p \in \mathbb{N}^*$ et $a_0, \dots, a_{p-1} \in \mathbb{K}$ tels que pour tout $i \in \mathbb{N}$,

$$u_{i+p} = \sum_{j=0}^{p-1} a_j u_{i+j}.$$

Dans ce cas, le polynôme $P = X^p - \sum_{j=0}^{p-1} a_j X^j$ vérifie $P(\delta)(u) = 0_{\mathcal{S}}$, donc l'ensemble

$$I(u) = \{P \in \mathbb{K}[X] \mid P(\delta)(u) = 0_{\mathcal{S}}\} = \ker(P \mapsto P(\delta)(u))$$

est un idéal non trivial de $\mathbb{K}[X]$. On appelle *polynôme minimal* de u le générateur unitaire $\pi(u)$ de $I(u)$ et *degré* de u le degré $d(u)$ de $\pi(u)$. Il est alors clair que le sous-espace vectoriel $\mathbb{K}[\delta](u) = \{P(\delta)(u) \mid P \in \mathbb{K}[X]\}$ de $\mathcal{S}$ admet pour base $\{\delta^i(u) \mid i = 0 \dots p-1\}$, et que la matrice de δ dans cette base est la matrice compagnon de $\pi(u)$:

$$\mathcal{C}_{\pi(u)} = \begin{pmatrix} 0 & \dots & \dots & 0 & a_0 \\ 1 & 0 & & \vdots & \vdots \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & a_{p-1} \end{pmatrix}$$

2.2 Polynôme minimal et premiers termes d'une suite récurrente linéaire scalaire

Par définition, une suite linéaire récurrente u de degré d est complètement déterminée par son polynôme minimal $\pi(u)$ et ses d premiers termes $u_0, \dots, u_{d-1}$. Dans ce qui suit, on s'intéresse au problème inverse : on se donne les $2d$ premiers termes d'une suite récurrente linéaire scalaire u de degré d, et on retrouve son polynôme minimal $\pi(u)$.

Soient $d \leq m$ deux entiers et u une suite récurrente linéaire scalaire de degré d dont on connaît les $2m$ premiers termes. On définit le polynôme $U = \sum_{i=0}^{2m-1} u_{2m-i-1} X^i$. Les polynômes de $I(u)$ de degré inférieur ou égal à m sont alors caractérisés par la proposition suivante :

Proposition 1. *Soit $P = \sum_{j=0}^{m} a_j X^j$ un polynôme de $\mathbb{K}[X]$ de degré inférieur ou égal à m. Les assertions suivantes sont équivalentes :*

(i) $P \in I(u)$, *ie.* $P(\delta)(u) = 0_{\mathcal{S}}$,

(ii) pour tout $i \leq m-1$, $\sum_{j=0}^{m} a_j u_{i+j} = 0$,

(iii) il existe $R \in \mathbb{K}[X]$ *tel que* $\deg(R) < m$ *et* $PU \equiv R \mod X^{2m}$.

L'implication *(i)⇒(ii)* est évidente, par définition de $I(u)$. Montrons que *(ii)⇒(i)* : il s'agit de montrer que si $\sum_{j=0}^{m} a_j u_{i+j} = 0$ pour tout $i \leq m-1$, alors ceci reste vrai pour tout $i \geq m$. Notons $\pi(u) = X^d - \sum_{k=0}^{d-1} b_k X^k$ le polynôme minimal de u. Pour un entier $p \geq m$, supposons que pour tout $i \leq p$, on ait $\sum_{j=0}^{m} a_j u_{i+j} = 0$. Alors

$$\sum_{j=0}^{m} a_j u_{p+1+j} = \sum_{j=0}^{m} a_j \left(\sum_{k=0}^{d-1} b_k u_{p+1+j-d+k} \right) = \sum_{k=0}^{d-1} b_k \left(\sum_{j=0}^{m} a_j u_{p+1-d+k+j} \right) = \sum_{k=0}^{d-1} b_k . 0 = 0.$$

Montrons maintenant que *(ii)⇔(iii)*. On a

$$PU = \sum_{j=0}^{m} a_j X^j \sum_{i=0}^{2m-1} u_{2m-i-1} X^i = \sum_{k=0}^{3m-1} X^k \sum_{\substack{i=0\dots 2m-1 \\ j=0\dots m \\ i+j=k}} a_j u_{2m-i-1}$$

Or pour tout $k \in \{m, \dots, 2m-1\}$, on a $0 \leq 2m-1-k \leq m-1$ donc si *(ii)* est vérifié,

$$\sum_{\substack{i=0\dots 2m-1 \\ j=0\dots m \\ i+j=k}} a_j u_{2m-i-1} = \sum_{j=0}^{m} a_j u_{2m-1-k+j} = 0.$$

On obtient donc $PU = R \mod X^{2m}$, où $R = \sum_{k=0}^{m-1} X^k \sum a_j u_{2m-i-1} + \sum_{k=2m}^{3m-1} X^{k-2m} \sum a_j u_{2m-i-1}$ est de degré strictement inférieur à m. Réciproquement, pour tout $0 \leq p \leq m-1$, $m \leq 2m-p-1 \leq 2m-1$ de sorte que si *(iii)* est vérifié,

$$\sum_{j=0}^{m} a_j u_{p+j} = \sum_{\substack{i=0\dots 2m-1 \\ j=0\dots m \\ i+j=2m-p-1}} a_j u_{2m-i-1} = 0,$$

et *(ii)* est vérifié.

2.3 Algorithme

Soit $m \in \mathbb{N}$, u une suite récurrente linéaire scalaire de degré inférieur ou égal à m dont on connaît les $2m$ premiers termes et $U = \sum_{i=0}^{2m-1} u_{2m-i-1} X^i$. D'après la proposition précédente, trouver les polynômes P de $I(u)$ de degré inférieur ou égal à m correspond à trouver les couples (P, R) avec $\deg(P) \leq m$, $\deg(R) < m$ et $PU = R \mod X^{2m}$. Un tel couple est donné par l'algorithme d'Euclide étendu :

POLYNÔME ANNULATEUR

$A \leftarrow X^{2m}$; $B \leftarrow U$; $C \leftarrow 0$; $D \leftarrow 1$;
while $\deg(B) < m$ **do**
  $(Q, R) \leftarrow$ (quotient, reste) de la division euclidienne de A par B
  $E \leftarrow C - QD$; $C \leftarrow D$; $D \leftarrow E$; $A \leftarrow B$; $B \leftarrow R$;
end while
return (D, B)

Pour s'assurer de la correction de cet algorithme, on montre par récurrence qu'à chaque étape, on a

$$\deg(C) \leq 2m - \deg(A), \quad \deg(D) \leq 2m - \deg(A), \quad m \leq \deg(B) \leq \deg(A),$$

$$CU \equiv A \mod X^{2m} \quad \text{et} \quad DU \equiv B \mod X^{2m}$$

(i) Initialisation : au commencement, on a
- $-\infty = \deg(C) \leq 0 = \deg(D) \leq 2m - \deg(A)$,
- $CU = 0 \equiv X^{2m} = A$ et $DU = U = B$.

(ii) Transmission : à chaque étape, on a
- $A = QB + R \Rightarrow R = A - QB \equiv CU - QDU = (C - QD)U = EU$,
- $\deg(Q) = \deg(A) - \deg(B)$ donc $\deg(QD) = \deg(A) - \deg(B) + \deg(D) \leq 2m - \deg(B)$ et $\deg(E) \leq \max(\deg(C), \deg(QD)) \leq \max(2m - \deg(A), 2m - \deg(B)) \leq 2m - \deg(B)$.

Lorsque l'algorithme s'arrête, on a donc $DU = B \mod X^{2m}$ avec $\deg(B) < m$ et $\deg(D) \leq 2m - \deg(A) \leq m$, ce qui prouve que l'algorithme renvoie bien un polynôme de $I(u)$ de degré inférieur ou égal à m.

3 Application à l'inversion de matrices creuses

3.1 Inverser une matrice creuse de $M_m(\mathbb{R})$ ou $M_m(\mathbb{C})$

On suppose ici que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit A une matrice inversible de $M_m(\mathbb{K})$, et $P = \sum_{i=0}^{d} a_i X^i$ un polynôme annulateur non trivial de A. On peut supposer que la valuation de P est nulle, ie. que $a_0 \neq 0$. On a alors

$$A^{-1} = -\frac{1}{a_0} \sum_{i=1}^{d} a_i A^{i-1},$$

de sorte que si l'on connaît un polynôme annulateur non trivial de A, on peut calculer l'inverse de A en n'utilisant que des multiplications par A. Pour appliquer cela à l'inversion d'une matrice creuse (ie. pour laquelle on privilégie l'opération de multiplication par A à toute autre opération matricielle - voir § 1), il suffit donc de savoir déterminer un polynôme annulateur non trivial de A.

Pour cela, l'idée de Wiedemann est de trouver une suite récurrente linéaire scalaire $u \in \mathcal{S}$ qui ait le même polynôme minimal que A. Il suffit ensuite d'appliquer l'algorithme d'Euclide étendu pour déterminer un polynôme annulateur non trivial de u, et donc de A.

Pour trouver cette suite récurrente linéaire scalaire, on choisit aléatoirement un vecteur $x \in \mathbb{K}^m$ et une forme linéaire $\lambda \in (\mathbb{K}^m)^*$, et on considère la suite u définie par $u_n = \lambda A^n x$. Le polynôme minimal de A annule clairement la suite u, de sorte que $\pi(u) | \pi_A$. Et un argument topologique permet de supposer en fait que $\pi(u) = \pi_A$. En effet,

(i) l'ensemble

$$\{\lambda \in (\mathbb{K}^m)^* \mid \pi\left((\lambda A^n)_{n\in\mathbb{N}}\right) \neq \pi_A\} = \bigcup_{\substack{P|\pi_A \\ P\neq\pi_A}} \ker(P(A^t))$$

est une union finie de sous-espaces vectoriels stricts de $(\mathbb{K}^m)^*$, donc est d'intérieur vide. Si on choisit $\lambda \in (\mathbb{K}^m)^*$ au hasard, on peut donc supposer que $\pi\left((\lambda A^n)_{n\in\mathbb{N}}\right) = \pi_A$.

(ii) pour tout $\lambda \in (\mathbb{K}^m)^*$ tel que $\pi\left((\lambda A^n)_{n\in\mathbb{N}}\right) = \pi_A$, l'ensemble

$$\{x\mathbb{K}^n \mid \pi\left((\lambda A^n x)_{n\in\mathbb{N}}\right) \neq \pi_A\} = \bigcup_{\substack{P|\pi_A \\ P\neq\pi_A}} \ker(\lambda P(A))$$

est une union finie de sous-espaces vectoriels stricts de $\mathbb{K}^m$, donc est d'intérieur vide. Si on choisit $x \in \mathbb{K}^n$ au hasard, on peut donc supposer que $\pi\left((\lambda A^n x)_{n\in\mathbb{N}}\right) = \pi_A$. On trouve donc directement notre suite.

3.2 Résoudre un système creux à coefficients dans $\mathbb{F}_q$

On suppose ici que $\mathbb{K} = \mathbb{F}_q$. Soit A une matrice creuse de $M_m(\mathbb{K})$ et $b \in \mathbb{K}^m$. On appelle *suite de Krylov* associée au système creux $Ax = b$ la suite $(A^n b)_{n\in\mathbb{N}} \in (\mathbb{K}^m)^{\mathbb{N}}$. Encore une fois, cette suite est récurrente linéaire, et si on en connaît un polynôme annulateur non trivial $P = \sum_{i=0}^{d} a_i X^i$ (que l'on peut supposer de valuation nulle), alors le vecteur

$$x = -\frac{1}{a_0} \sum_{i=1}^{d} a_i A^{i-1} b.$$

est une solution du système $Ax = b$.

Comme précédemment, la connaissance d'un polynôme annulateur non trivial de la suite de Krylov associé à un système creux $Ax = b$ permet donc de résoudre ce système en n'utilisant que des multiplications par A. Pour trouver un

tel polynôme, on choisit aléatoirement une forme linéaire $\lambda \in (\mathbb{K}^m)^*$, et on considère la suite u définie par $u_n = \lambda A^n x$. Le polynôme minimal de la suite de Krylov $(A^n b)_{n\in\mathbb{N}}$ annule clairement la suite u, de sorte que $\pi(u)|\pi_{(A^n b)_{n\in\mathbb{N}}}$. On a alors deux possibilités :

(i) si $\pi(u) = \pi_{(A^n b)_{n\in\mathbb{N}}}$, ie. si $\pi(u)(A)b = 0$, on a terminé : on peut résoudre le système.

(ii) sinon, $b' = \pi(u)(A)b \neq 0$. La suite de Krylov associée au système $Ax = b'$ a pour polynôme minimal $\pi_{(A^n b')_{n\in\mathbb{N}}} = \pi_{(A^n b)_{n\in\mathbb{N}}}/\pi(u)$, de degré inférieur ou égal à $n - \deg(\pi(u))$. On choisit alors au hasard une nouvelle forme linéaire μ, et on calcule le polynôme minimal de la suite récurrente linéaire scalaire u' définie par $u'_n = \mu A^n b'$, qui a son tour est soit $\pi_{(A^n b')_{n\in\mathbb{N}}}$, soit un diviseur strict de $\pi_{(A^n b')_{n\in\mathbb{N}}}$, et ainsi de suite. On finit par obtenir le polynôme minimal de $(A^n b)_{n\in\mathbb{N}}$.

L'algorithme peut donc s'écrire de la manière suivante :

ALGORITHME DE WIEDEMANN

$x \in \mathbb{K}^m$ choisi au hasard; $\quad P \leftarrow 1$; $\quad d \leftarrow 0$;
while $P(A) \neq 0$ **do**
  $\lambda \in (\mathbb{K}^m)^*$ choisi au hasard ;
  calcul des $2(n-d)$ premiers termes de la suite $(\lambda A^n x)_{n\in\mathbb{N}}$;
  $\pi \leftarrow$ le polynôme minimal de cette suite ;
  $x \leftarrow \pi(A)x$; $\quad P \leftarrow P\pi$; $\quad d \leftarrow d + \deg(\pi)$;
end while
return P ;

Il reste alors à montrer que le nombre d'étapes dans cet algorithme reste petit. Dans *Solving sparse linear equations over finite fields* (*IEEE trans. inf. theory* 32, 1986), Wiedemann montre que la probabilité de ne pas avoir obtenu $\pi_{(A^n b)_{n\in\mathbb{N}}}$ après k étapes est inférieure à

$$\log\left(\frac{q^k}{q^k - 1}\right).$$

En particulier, lorsque q est grand, la probabilité de ne pas avoir obtenu $\pi_{(A^n b)_{n\in\mathbb{N}}}$ du premier coup est proche de 0.

4 Questions et remarques

4.1 Questions

On pourra traiter les problèmes suivants :

1. sur les suites récurrentes linéaires :
   (a) soit $P = \prod_{i=1}^{\ell}(X - \lambda_i)^{\mu_i}$ un polynôme de $\mathbb{C}[X]$. Donner une base de l'espace vectoriel des suites récurrentes linéaires annulées par P.
   (b) présenter une méthode générale adoptant un point de vue matriciel pour trouver une base de l'espace vectoriel des suites récurrentes linéaires annulées par un polynôme P.
   (c) donner des exemples où apparaissent des suites récurrentes linéaires. On pourra par exemple exposer rapidement :
   – le calcul du déterminant de la matrice

$$\begin{pmatrix} a & b & 0 & \dots & 0 \\ c & a & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & a & b \\ 0 & \dots & 0 & c & a \end{pmatrix}$$

   – le nombre de compositions d'un entier n dont les sommants sont dans un ensemble fini $U = \{u_1, \dots, u_p\} \subset \mathbb{N}^*$, ie.

$$\mathcal{C}_n^U = \operatorname{card}\left(\bigcup_{\ell\in\mathbb{N}^*} \{(x_1, \dots, x_\ell) \in U \mid x_1 + \dots + x_\ell = n\}\right).$$

(d) pourquoi ne peut-on pas retrouver le polynôme minimal d'une suite récurrente linéaire u de degré d avec moins que $2d$ termes ?

(e) soit $u = (u_n)_{n\in\mathbb{N}}$ une suite récurrente linéaire de degré d. Soit $p \in \mathbb{N}$. Montrer que la suite v définie par $v_n = u_n \mod p$ est périodique. Donner une borne pour sa période.

2. sur la complexité de l'algorithme :
   (a) quels algorithmes utilise-t-on en général pour inverser une matrice ? Quelle est leur complexité ?
   (b) discuter la complexité de l'algorithme.
   (c) cet algorithme est-il efficace lorsque la matrice n'est pas creuse.
3. programmation :
   (a) implémenter l'algorithme sur $\mathbb{R}$.
   (b) développer un système de calcul sur un corps fini (pas nécessairement premier !) et implémenter l'algorithme de Wiedemann. Étudier le nombre d'étapes nécessaires à l'obtention du polynôme minimal de la suite de Krylov d'un système creux.

4.2 Remarques et références

L'algèbre linéaire creuse est une situation particulière qui se rencontre en particulier dans le cadre des algorithmes de logarithme discret et de factorisation d'entiers (crible quadratique). On trouvera l'algorithme de Wiedemann dans *Modern Computer Algebra* de J. Von Zur Gathen & J. Gerhard ou dans *Méthodes matricielles, Introduction à la complexité algébrique* de J. Abdeljaoued, H. Lombardi.

Il me semble que l'étude de cet algorithme est une bonne occasion pour revoir les algorithmes usuels d'inversion de matrice d'une part, et pour programmer dans un corps fini d'autre part. Par ailleurs, il peut constituer un bon exemple dans les leçons d'agrégation suivantes :

- Résolution d'un système d'équations linéaires
  On pourra discuter des différentes méthodes de résolution d'un système linéaire et du gain en complexité lorsque ce système est creux.
- Polynômes d'endomorphismes, polynômes annulateurs
  On utilise des polynômes annulateurs pour inverser une matrice. On présentera en particulier la discussion du paragraphe 3.1
- Suites réelles ou vectorielles définies par itération