COURS 28

Algèbre linéaire creuse : algorithme de Wiedemann

1. Introduction

Dans cette section, on aborde des questions radicalement différentes : on ne cherche plus à effectuer les produit, inversion, ... de matrices quelconques, mais à manipuler des matrices *creuses*.

On ne commencera pas par définir précisément de ce qu'est une matrice creuse. L'approche consiste à donner des algorithmes dont la complexité s'exprime en fonction du nombre d'entrées non nulles des matrices en entrées. De la complexité des algorithmes « creux » va découler une borne, typiquement de l'ordre $O(n)$ pour des matrices de taille $n \times n$, sur le nombre de coefficients non nuls pour que le changement de modèle soit pertinent.

Dans tout ce qui suit, on considère donc une matrice A de taille $n \times n$ et vérifiant les hypothèses suivantes :

- A est inversible,
- A contient s entrées non nulles.

On va montrer un algorithme dû à Wiedemann qui permet de calculer l'unique solution du système $Ax = y$ avec une complexité en $O(ns)$; l'algorithme original est plus général que celui présenté ici, en ce qu'il permet de traiter les matrices rectangulaires ou carrées et non inversibles, mais il se ramène au cas carré inversible.

Remarquons que si s est de l'ordre de n^2, caractérisant donc des matrices plutôt pleines, on retombe dans une complexité de l'ordre de $O(n^3)$. Le cas intéressant est celui où s est de l'ordre de n, auquel cas l'algorithme est quadratique en n : on gardera en tête que l'exposant de l'algèbre linéaire creuse est 2 dans les bons cas.

1.1. Polynôme minimal et résolution de systèmes.

L'algorithme de Wiedemann passe par le calcul du *polynôme minimal* de A. Rappelons sa définition.

Si k est le corps de base, il est possible d'associer à tout polynôme en une variable $P \in k[T]$ la matrice $P(A)$. On sait que d'après le théorème de Cayley-Hamilton, $\chi(A) = 0$, où χ est le polynôme caractéristique de A. Le polynôme minimal de A est le polynôme unitaire de plus petit degré ayant cette propriété (on montre facilement que cette propriété le rend unique).

Soit P le polynôme minimal de A, supposé connu. Puisqu'on a supposé A inversible, le terme constant de P n'est pas nul (car il divise le terme constant du polynôme caractéristique, qui n'est lui-même pas nul).

L'égalité $P(A) = 0$ se réécrit sous la forme

$$A^{-1} = \frac{-1}{p_0} \left(A^{m-1} + p_{m-1} A^{m-2} + \cdots + p_1 I_n \right)$$

Pour résoudre le système $Ax = y$, on va calculer $x = A^{-1}y$, c'est-à-dire

$$\frac{-1}{p_0}\left(A^{m-1}y + p_{m-1}A^{m-2}y + \cdots + p_1 y\right).$$

Ainsi, il suffit de calculer les itérés $y, Ay, \ldots, A^{m-1}y$, puis d'additionner les termes $p_1 y, p_2 Ay, \ldots, A^{m-1}y$, pour retrouver x.

- Chacun des $A^i y$ s'obtient à partir de $A^{i-1}y$ en $O(s)$ opérations.
- Chacun des produits par p_i, et chaque addition, se fait en $O(n)$ opérations. Remarquer que $n \leq s$ (hypothèse d'inversibilité de A).

Au total, on a donc $O(ns)$ opérations à faire pour résoudre le système si on connaît le polynôme minimal de A.

1.2. Calcul du polynôme minimal. écrivons le polynôme minimal de A sous la forme

$$P = T^m + p_{m-1}T^{m-1} + \cdots + p_0.$$

Alors la suite des puissances A satisfait la récurrence linéaire

$$A^{k+m} + p_{m-1}A^{k+m-1} + \cdots + p_0 A^k = 0,$$

et ne satisfait pas de récurrence d'ordre plus petit.

Pour tous vecteurs u et v, la suite (de nombres) $N_i := u^t A^i v$ vérifie donc

$$N_{k+m} + p_{m-1}N_{k+m-1} + \cdots + p_0 N_k = 0.$$

Si u et v sont mal choisis (nuls, par exemple), la suite N_i satisfait une récurrence d'ordre plus petit que m. La proposition suivante montre qu'en choisissant u et v au hasard, on tombe vraisemblablement sur une suite N_i qui ne satisfait pas de récurrence d'ordre plus petit.

PROPOSITION 1. *Il existe un polynôme non nul D dans $k[U_1, \ldots, U_n, V_1, \ldots, V_n]q$, de degré au plus $2n$, tel que si $D(u, v)$ est non nul, la suite N_i associée à u et v ne satisfait pas de récurrence d'ordre plus petit que m.*

L'idée est de choisir u et v au hasard. Si on n'a pas de chance, $D(u, v)$ est nul ; sinon, la suite des N_i associée à u et v permet de retrouver P par un calcul d'approximants de Padé (attention, on ne connaît pas explicitement le polynôme D, mais son existence assure que la plupart des choix de u et v sont "chanceux").

Algorithme de Wiedemann.:

Entrée :: la matrice A.

Sortie :: son polynôme minimal.

1. Choisir des vecteurs u et v aléatoires.
2. Pour $0 \leq i \leq 2n$, calculer les vecteurs $A^i v$, puis les scalaires $N_i = u^t A^i v$.
3. Retrouver la récurrence satisfaite par les N_i.

La complexité de l'étape 2 est $O(n)$ produits matrice-vecteur ou vecteur-vecteur, chacun prenant au pire $O(s)$ opérations ; au total on obtient donc $O(ns)$ opérations pour cette phase. L'algorithme d'approximants de Padé est négligeable, puisqu'il ne demande que $O(\mathsf{M}(n)\log(n)) \leq sn$ opérations.

Bibliographie