Forme de Frobenius et vecteurs cycliques

Daniel AUGOT[*], Paul CAMION[†]

Abstract

Le premier algorithme présenté, de complexité $O(n^3)$, produit un vecteur cyclique d'une matrice de $M_n(\mathbf{k})$ dont le polynôme caractéristique est sans facteurs multiples. Ceci fournit un algorithme *déterministe* de calcul d'une base normale de $\mathbb{F}_{q^n}$ de complexité $O(n^3)$, sur la donnée d'une *représentation* de l'opérateur de Frobenius de $\mathbb{F}_{q^n}$ sur $\mathbb{F}_q$. Le second est un algorithme de calcul de la forme de rationnelle canonique d'une matrice quelconque. Il est de complexité $O(n^3 m_{\mathbf{A}})$ où $m_{\mathbf{A}}$ est le nombre de facteurs du polynôme caractéristique de A. Cette complexité est en moyenne $O(n^3 \log n)$ pour $\mathbf{k} = \mathbb{F}_q$. La connaissance effective du centralisateur d'une matrice est obtenue.

Frobenius Form and cyclic vectors

Abstract

The first presented algorithm, whose complexity is $O(n^3)$, produces a cyclic vector for a matrix of $M_n(\mathbf{k})$ whose characteristic polynomial is square-free. This gives a *deterministic* algorithm of complexity $O(n^3)$ for computing a normal basis of $\mathbb{F}_{q^n}$, given a *presentation* of $\mathbb{F}_{q^n}$ and a matrix representing the Frobenius operator of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. The second algorithm computes the rational canonical form of any matrix. Its complexity is $O(n^3 m_{\mathbf{A}})$ where $m_{\mathbf{A}}$ is the number of factors of the characteristic polynomial of A. The average complexity is $O(n^3 \log n)$ for $\mathbf{k} = \mathbb{F}_q$. Moreover, the real knowledge of the centralizer of a matrix is obtained.

[*]Paris 6, INRIA Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex - France

[†]CNRS, INRIA

Abridged English version – In the first part of this paper, we show how to compute a cyclic vector for a matrix whose characteristic polynomial is square-free. Let $\mathbf{A} \in M_n(\mathbf{k})$ be given. A cyclic vector for $\mathbf{A}$ is a vector v such that the minimal polynomial of v equals the minimal polynomial of $\mathbf{A}$. Shift-Hessenberg matrices are introduced in definition 1. are used through all the paper. We have that for every matrix $\mathbf{A}$ in $M_n(\mathbf{k})$, there exists a Shift-Hessenberg matrix $\mathbf{H}$ and an invertible matrix $\mathbf{P}$ such that $\mathbf{H} = \mathbf{PAP^{-1}}$. The matrices $\mathbf{H}$ and $\mathbf{P}$ can be obtained in $O(n^3)$ elementary operations.

In the first section, we shall assume that the characteristic polynomial of $\mathbf{A}$ is square-free. Lemma 1 is the basic tool for constructing a cyclic vector, and shows how to compute a cyclic vector for a matrix which is block-triangular with two blocks, assuming that a cyclic vector is known for each block.

We then describe the strategy of the algorithm. In order to compute a cyclic vector for the matrix $\mathbf{A}$, a Shift-Hessenberg form $\mathbf{H}$ of $\mathbf{A}$ is first computed. Then a matrix

$$\mathbf{H}_{split} = \begin{bmatrix} \mathbf{H}'_{B_I,B_I} & \mathbf{H}'_{B_I,B_J} \\ 0 & \mathbf{H}'_{B_J,B_J} \end{bmatrix}$$

is obtained, such that $\mathbf{H}_{split}$ is similar to $\mathbf{H}$, and: either both blocks have size $\leq 2n/3$, or the first block has size greater than $2n/3$ and is a companion matrix. The algorithm is recursively applied on both matrices, in order to find v_{B_I}, v_{B_J} which are cyclic vectors for $\mathbf{H}'_{B_I,B_I}$ respectively. Companion matrices are dealt with on terminal cases. From lemma 1, a cyclic vector v for H_{split} can be constructed. Keeping track of all changes of bases, a cyclic vector for $\mathbf{A}$ is obtained. It is shown that the complexity for this whole procedure is $O(n^3)$ (proposition 1). In particular, a normal basis for $\mathbb{F}_{q^n}$ can be computed in $O(n^3)$ elementary operations, on the data of a matrix representing the Frobenius automorphism (corollary 1).

In the second section, we are concerned with the problem of finding the Frobenius form of a matrix $\mathbf{A}$. The first step is to compute bases for characteristic subspaces. For $\mathbf{A} \in M_n(\mathbf{k})$, let $m_{\mathbf{A}}$ be the number of irreducible factors of the characteristic polynomial of $\mathbf{A}$. In the case where $\mathbf{k}$ is a finite field, relying on the results of Stong [1], we show that the expected value of $m_{\mathbf{A}}$ is $O(\log n)$, when n goes to infinity. Then, if the factorization of the characteristic polynomial of $\mathbf{A}$ is known, we can compute a basis for each characteristic subspace V_i, and the restriction of $\mathbf{A}$ to the V_i's in $O(n^3 m_{\mathbf{A}})$ elementary operations (lemma 2).

Once the characteristic subspaces of matrix $\mathbf{A}$ have been computed, we compute the Frobenius form of the restriction of matrix $\mathbf{A}$ to characteristic subspaces. Let $\mathbf{A}$ be a matrix whose characteristic polynomial is $p(X)^m$, $p(X)$ being irreducible. A Shift-Hessenberg form $\mathbf{H}$ of $\mathbf{A}$ is computed. If $\mathbf{H}$ is a companion matrix, the result is $\mathbf{H}$. We show how to proceed: either for augmenting the size of the first block, either for finding a basis for an invariant subspace supplementary to the one determined by the first block. This process ends in $O(n^3m)$ elementary steps. Resting on that basic algorithm, we finally obtain:

Theorem 1 *If the factorization of the characteristic polynomial of $\mathbf{A}$ is known, it is possible to compute the rational canonical form and the matrix for changing basis in $O(n^3 m_{\mathbf{A}})$ elementary operations, i.e. an asymptotic average complexity of $O(n^3 \log n)$ operations.*

1 Calcul d'un vecteur cyclique

Soit $\mathbf{k}$ un corps commutatif, et soit $\mathbf{A} \in M_n(\mathbf{k})$. Soit $v \in \mathbf{k}^n$; le *polynôme minimal* de v est le polynôme unitaire de plus petit degré tel que $p(\mathbf{A})v = 0$. Un vecteur v est *cyclique* si le polynôme minimal de v est égal au polynôme minimal de $\mathbf{A}$.

Dans cette partie, nous supposons donnée une matrice $\mathbf{A} \in M_n(\mathbf{k})$ dont le polynôme caractéristique est sans facteurs multiples. Dans ce cas, le polynôme caractéristique de $\mathbf{A}$ est égal au polynôme minimal de $\mathbf{A}$, et nous présentons un algorithme de construction d'un vecteur cyclique.

1.1 Forme de Hessenberg à décalage

Définition 1 *Une matrice $\mathbf{H}$ de $M_n(\mathbf{k})$ est de Hessenberg à décalage si elle est de la forme*

$$\mathbf{H} = \begin{bmatrix} 0 & & & \times & & & \times \\ 1 & & & \times & & & \times \\ & \ddots & & \times & & & \\ & & 1 & \times & & & \times \\ & & & 0 & & & \times \\ & & & & 1 & & \times \\ & & & & & \ddots & \end{bmatrix}$$

c'est-à-dire

$$\begin{array}{lcl}(h_{i+1,i} \neq 0) & \Rightarrow & (h_{i+1,i} = 1 \ \textit{et}\ \forall j \neq i+1 \ h_{j,i} = 0). \\ (h_{i+1,i} = 0) & \Rightarrow & (\forall j > i+1 \ h_{j,i} = 0)\end{array}$$

Le paramètre m d'une telle matrice est le nombre de zéros dans la première sous-diagonale augmenté de un. C'est aussi le nombre de matrices compagnons qui forment ses blocs diagonaux. Cette forme, cas particulier de la forme de Hessenberg [2], a déjà été utilisée dans [3].

Notation 1 *Une matrice de Hessenberg à décalage* $\mathbf{H}$ *détermine une famille croissante* $V_1, V_2, \ldots, V_m$ *d'espaces invariants sous* $\mathbf{H}$. *Soit* ϵ_{i+1} *le* $(t_i + 1)$*-ième vecteur unité où* $t_0 = 0$ *et où* t_i *est la somme des tailles des* i *premiers blocs diagonaux de* $\mathbf{H}$. *De cette manière* $V_1 = Vect(\epsilon_1, H\epsilon_1, \ldots)$, $V_2 = Vect(\epsilon_1, H\epsilon_1, \ldots, \epsilon_2, H\epsilon_2, \ldots)$, $\ldots$, $V_m = Vect(\epsilon_1, H\epsilon_1, \ldots, \epsilon_2, H\epsilon_2, \ldots, \epsilon_m, H\epsilon_m, \ldots)$.

Théorème 1 *Toute matrice est semblable à une matrice de Hessenberg à décalage. La forme de Hessenberg à décalage et la matrice de passage peuvent être calculées en* $O(n^3)$.

Preuve : Il s'agit de faire une élimination de Gauss en choisissant les pivots sur la première sous-diagonale. Des permutations de lignes et de colonnes obtenues par conjugaison sont effectuées pour obtenir un pivot non nul sous la première sous-diagonale.

□

Si le paramètre m est égal à 1, la matrice de Hessenberg est une matrice compagnon, et le vecteur ${}^t(1, 0, \ldots, 0)$ est un vecteur cyclique. Le lemme suivant traite le cas où une matrice générale $\mathbf{A}$ a deux blocs diagonaux. Pour une matrice ayant plusieurs blocs diagonaux, nous notons B_i l'ensemble d'indices correspondant au i-ième bloc, et pour chaque B_i, pour tout vecteur $v \in \mathbf{k}^n$, nous notons v_{B_i} la projection de v dans $\mathbf{k}^{B_i}$. Pour tout vecteur v_{B_i} dans $\mathbf{k}^{B_i}$ nous notons $v^\star_{B_i}$ le vecteur de $\mathbf{k}^n$ tel que $v^\star_{B_i\, j} = 0$ si $j \notin B_i$, et $v^\star_{B_i\, j} = v_{B_i\, j}$ si $j \in B_i$.

Lemme 1 *Soit* $\mathbf{A}$ *une matrice donnée par blocs*

$$\begin{bmatrix} \mathbf{A}_{B_1,B_1} & \mathbf{A}_{B_1,B_2} \\ 0 & \mathbf{A}_{B_2,B_2} \end{bmatrix}$$

et soient v_{B_1}, v_{B_2} deux vecteurs cycliques pour $\mathbf{A}_{B_1,B_1}$ et $\mathbf{A}_{B_2,B_2}$ respectivement et de polynômes minimaux respectifs $f_1(X)$ et $f_2(X)$. Si $f_1(X)$ et $f_2(X)$ sont premiers entre eux, un vecteur cyclique pour $\mathbf{A}$ peut être calculé en $O(n^3)$.

Preuve : Soit $u = (u_{B_1}, u_{B_2})$ vérifiant

$$\begin{aligned} u_{B_2} &= v_{B_2} && (1) \\ f_2(\mathbf{A}_{B_1,B_1})u_{B_1} + (f_2(\mathbf{A})u^{\star}_{B_2})_{B_1} &= v_{B_1}. && (2) \end{aligned}$$

Un tel vector u existe car $f_2(\mathbf{A}_{B_1,B_1})$ est inversible. On vérifie que son polynôme minimal est $f_1(X)f_2(X)$. Notons que résoudre le système (refeq:2) ne nécessite pas de calculer $f_2(\mathbf{A})$. En effet $f_2(X)$ est inversible modulo $f_1(X)$, d'inverse $g_2(X)$, et résoudre 2 revient à appliquer $g_2(\mathbf{A})$ à un vecteur. $\square$

Observons que la forme de Frobenius obtenue au paragraphe 2.2 fournit un vecteur cyclique pour tout opérateur linéaire.

1.2 L'algorithme

Nous en donnons ici une description succinte. On donne une matrice $\mathbf{A} \in M_n(\mathbf{k})$ dont le polynôme caractéristique est sans facteurs multiples. On obtient une matrice de passage $\mathbf{P}$ vérifiant $\mathbf{A} = \mathbf{PCP}^{-1}$, $\mathbf{C}$ étant une matrice compagnon.

Etape 1 : Calcul d'une forme de Hessenberg à décalage $\mathbf{H}$ de la matrice $\mathbf{A}$. Si $\mathbf{H}$ est une matrice compagnon, l'algorithme s'arrête.

Etape 2 (partitionnement) : Celui-ci produit une matrice de Hessenberg à décalage $\mathbf{H_{I,J}}$ semblable à $\mathbf{H}$,

$$\mathbf{H_{I,J}} = \begin{bmatrix} \mathbf{H}'_{B_I,B_I} & \mathbf{H}'_{B_I,B_J} \\ 0 & \mathbf{H}'_{B_J,B_J} \end{bmatrix}$$

où l'on note B_I (resp.B_J) l'ensemble $\cup_{i\epsilon I} B_i$ (resp.$\cup_{i\epsilon J} B_i$) et I et J forment une partition de $[m]$. Elle est telle que : ou bien les deux blocs diagonaux, qui sont eux-mêmes des formes de Hessenberg à décalage ont une taille d'au plus $2n/3$, ou bien le bloc $\mathbf{H}'_{B_I,B_I}$ est une matrice compagnon de taille supérieure à $2n/3$.

Etape 3 (récurrence) : L'algorithme est appliqué récursivement à partir de l'étape 2 soit à $\mathbf{H}'_{B_I,B_I}$ et $\mathbf{H}'_{B_J,B_J}$, soit à $\mathbf{H}'_{B_J,B_J}$ seule.

Etape 4 (reconstruction) : En utilisant le lemme 1, on reconstruit un vecteur cyclique pour $\mathbf{k}^n$.

Sortie : Par multiplication des matrices de changements de base successifs, on obtient $\mathbf{P}$.

Remarque Notons que l'étape 2 comporte éventuellement une et alors une seule permutation de blocs diagonaux obtenue par conjugaison suivie d'une remise en forme de Hessenberg. L'aboutissement de cette étape résulte alors d'un lemme.

Voici le résultat que l'on peut alors démonter.

Proposition 1 *Un vecteur cyclique pour* $\mathbf{A}$ *peut être construit en* $O(n^3)$ *opérations élémentaires.*

Considérons le corps $\mathbf{K} = \mathbb{F}_{q^n}$, on appelle base normale de $\mathbf{K}$ une famille $\gamma, \gamma^q, \ldots, \gamma^{q^{n-1}}$ qui engendre $\mathbf{K}$ par combinaisons $\mathbb{F}_q$-linéaires. Donc $\gamma, \gamma^q, \ldots, \gamma^{q^{n-1}}$ est une base normale de $\mathbb{F}_{q^n}$ si et seulement si γ est un vecteur cyclique pour l'automorphisme de Frobenius $\pi : x \mapsto x^q$, vu comme endomorphisme du $\mathbf{k}$-espace vectoriel $\mathbf{k}^n$, où $\mathbf{k} = \mathbb{F}_q$. Le polynôme minimal de π est $X^n - 1$ et si $\gcd(n, q) = 1$, $X^n - 1$ est sans facteurs multiples. Notre algorithme s'applique directement dans ce cas là. Dans le cas général nous écrivons $n = p^e n_1$, avec $\gcd(n_1, p) = 1$. Notre algorithme permet de calculer une base normale pour $\mathbb{F}_{q^{n_1}}$. Pour $\mathbb{F}_{q^{p^e}}$, on constatera que le dernier vecteur ϵ_m déterminé par une forme de Hessenberg à décalage représentant l'opérateur de Frobenius, est cyclique. On sait alors que le produit de deux éléments cycliques respectivement pour l'opérateur de Frobenius de $\mathbb{F}_{q^{n_1}}$ et de $\mathbb{F}_{q^{p^e}}$ est cyclique pour celui de $\mathbb{F}_{q^{n_1 p^e}}$ (Par exemple [4]).

Corollaire 1 *Une base normale de* $\mathbb{F}_{q^n}$*, n quelconque, peut être calculée en* $O(n^3 \log q)$ *opérations sur* $\mathbb{F}_q$*.*

Le coût supplémentaire $O(n^3 \log q)$ provient du précalcul de la matrice de l'endomorphisme de Frobenius.

2 Calcul de la forme de Frobenius

2.1 Algorithme pour le calcul des sous-espaces caractéristiques

Pour toute matrice $\mathbf{A}$ de $M_n(q)$ soit m_A le nombre de facteurs irréductibles du polynôme caractéristique de $\mathbf{A}$, comptés avec multiplicités. D'après R. Stong

[1], la valeur moyenne de m_A est asymptotiquement $\log n$ pour une matrice inversible à coefficients dans $\mathbb{F}_q$. Notons $C_{\mathbf{T}}(X)$ le polynôme caractéristique d'une matrice $\mathbf{T}$. Le résultat de R. Stong peut être prolongé de $GL(n, q)$ à $M_n(q)$ en développant l'argument suivant. On montre d'abord que si $\mathbf{A}$ est la somme directe des matrices $\mathbf{B}$ et $\mathbf{C}$, où $\gcd(C_{\mathbf{B}}(X), C_{\mathbf{C}}(X)) = 1$, alors le centralisateur $\mathcal{Z}(\mathbf{A})$ est le produit direct de $\mathcal{Z}(\mathbf{B})$ et de $\mathcal{Z}(\mathbf{C})$. Partant de la forme rationnelle canonique $\mathbf{F}$ d'une matrice $\mathbf{A}$ de $M_n(q)$, on l'écrit comme la somme directe d'une matrice nilpotente $\mathbf{F}_1$ et d'une matrice inversible $\mathbf{F}_2$. On applique alors le théorème de R. Stong pour vérifier que l'espérance du nombre de facteurs irréductibles de $\mathbf{F}_1 + \mathbf{I}$ (respectivement de $\mathbf{F}_2$), donc de $\mathbf{A}$ est asymptotiquement $\log n$. Ce développement conduit à un algorithme de calcul du centralisateur d'une matrice sur un corps et à son énumération pour un cors fini.

Le polynôme caractéristique $C_{\mathbf{A}}(X)$ peut être calculé en $O(n^3)$ opérations élémentaires, sur la donnée d'une forme de Hessenberg [2]. Sur $\mathbb{F}_q$ la factorisation de $C_{\mathbf{A}}(X)$ n'est pas coûteuse : $O(n^3 \log q)$ opérations dans $\mathbb{F}_q$ en myenne, par algorithme probabiliste.

Soit $C_{\mathbf{A}}(X) = P_1(X)^{r_1} \cdots P_k(X)^{r_k}$. On sait que les sous-espaces caractéristiques sont les $V_i = \ker P_i^{r_i}(\mathbf{A})$. On obtient le résultat que voici.

Lemme 2 *Etant donnée la factorisation de $C_{\mathbf{A}}(X)$, les sous-espaces caractéristiques V_i et la restriction de $\mathbf{A}$ aux V_i peuvent être calculés en $O(n^3 m_{\mathbf{A}})$ opérations élémentaires dans le corps $\mathbf{k}$. Pour $\mathbf{k} = \mathbb{F}_q$, la complexité est en moyenne de $O(n^3 \log n)$, la complexité maximale est $O(n^{3.5})$ pour tout corps.*

2.2 Algorithme pour le calcul de la forme de Frobenius sur un sous-espace caractéristique

Nous supposons le calcul précédent déjà accompli, et nous recherchons la forme rationnelle canonique de la restriction de $\mathbf{A}$ à un sous-espace caractéristique. Soit donc $\mathbf{A}$ une matrice de polynôme caractéristique $p(X)^m$, $p(X)$ irréductible.

Nous utilisons la notation qv pour $q(\mathbf{A})v$, $v \in \mathbf{k}^n$, $q \in \mathbf{k}[X]$. Voici une description succinte de cet algorithme.

On calcule une forme $\mathbf{H}$ de Hessenberg à décalage de $\mathbf{A}$. Si celle-ci ne contient qu'un bloc, $\mathbf{H}$ est la matrice cherchée.

Sinon, nous cherchons à remplacer chaque ϵ_i, $i > 1$ par $\epsilon'_i = \epsilon_i + a_i \epsilon_1$, $a_i \in \mathbf{k}[X]$, pour obtenir une nouvelle base dans laquelle $\mathbf{A}$ sera représentée par une matrice de Hessenberg à décalage, notée à nouveau $\mathbf{H}$, telle que $\forall\, k, i > 1, (\mathbf{H}^k \epsilon'_i)_{B_1} = 0$. Deux cas se présentent.

Pour chaque $i > 1$ il existe $a_i \in \mathbf{k}[X]$ tel que $\forall\, k$, $(\mathbf{H}^k \epsilon'_i)_{B_1} = 0$. Les vecteurs $\mathbf{H}^j \epsilon'_i$, $i > 1$, $j \geq 0$, sous-tendent un supplémentaire de V_1, invariant par $\mathbf{H}$. On applique à nouveau l'algorithme au le deuxième bloc diagonal de la matrice qui représente $\mathbf{A}$ dans la base $\epsilon_1, \epsilon'_2, \ldots, \epsilon'_m$, le premier bloc diagonal étant une matrice compagnon. Dans le cas contraire, il existe un ϵ_i pour lequel $(\mathbf{H}^k(\epsilon_i + a_i \epsilon_1))_{B_1} \neq 0$ pour tout $a_i \in \mathbf{k}[X]$. Par une conjugaison faite de permutations adéquates de lignes et de colonnes on obtient une matrice relative à une nouvelle base où ce vecteur ϵ_i est placé en première position et il en résultera que l'algorithme de réduction en forme de Hessenberg à décalage fournit une matrice dont le premier bloc est de taille accrue.

De cette manière, ou bien on augmente la taille du premier bloc, ou bien la matrice obtenue sera la somme directe du premier bloc et d'une matrice de Hessenberg à décalage. Ceci nous assure de l'aboutissement de l'algorithme.

La taille du premier bloc augmente d'un multiple de $\deg p$, il en résulte que le coût de calcul de la forme rationnelle canonique d'une telle matrice est dans le pire cas, $O(n^3 m)$. D'où le théorème

Théorème 2 *La factorisation du polynôme caractéristique d'une matrice* $\mathbf{A}$ *étant donnée, sa forme rationelle canonique et la matrice de passage peuvent être calculées en* $O(n^3 m_{\mathbf{A}})$*, où* $m_{\mathbf{A}}$ *est le nombre de facteurs irréductibles du polynôme caractéristique de* $\mathbf{A}$*, comptés avec multiplicité. Lorsque* $\mathbf{k} = \mathbb{F}_q$*, la complexité de l'algorithme décrit est en moyenne de* $O(n^3 \log n)$ *opérations élémentaires dans* $\mathbf{k}$*.*

"Résumé d'une texte qui sera conservé 5 ans par les Archives de l'Académie et dont copie peut être obtenue de l'Académie."

References

[1] Richard Stong *Some Asymptotic Results on Finite Vector Spaces*, Advances in Applied Mathematics, **9**, pp. 167-199 (1988).

[2] J. H. Wilkinson *The Algebraic Eigenvalue Problem* Clarendon Press. Oxford. First published 1965. Reprinted 1992.

[3] Patrick Ozello *Calcul exact des formes de Jordan et de Frobenius d'une matrice*, Thèse de 3ème cycle, Université Scientifique Technologique et Médicale de Grenoble, 1987.

[4] I.F. Blake, X.H. Gao, R.C. Mullin, S.A.Vanstone, T. Yaghoobian *Applications of finite fields* Kluwer Academic Publishers, Boston, Dordrecht, London. 1993.