Normal Bases and Canonical Rational Form (Over Finite Fields)

Daniel Augot
INRIA – Projet CODES

January 23, 1995

[summary by François Morain]

1. Introduction

Let $k = \mathbb{F}_q$ be the finite field with q elements (q a prime power p^r, r any nonnegative integer). For the basic properties of finite fields, as well as an introduction to normal bases, etc., we urge the unfamiliar reader to read [3, 4, 5].

Let A be a k-linear operator of k^n and call M the associated matrix, that is an element of $\mathcal{M}_n(k)$. The aim of this talk is to introduce the so-called *Shift-Hessenberg form* of M (SHS form for short) and describe its properties. In particular, it will be shown that there exists a fast algorithm for computing H.

Having the SHS form of M enables us to solve several problems. First, we can find cyclic vectors for A and therefore find a normal basis for $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. We can also find the minimal polynomial of M. Moreover, if we have the factorisation of the characteristic polynomial of M, we can compute the characteristic subspaces of A and get the *Frobenius form* (a.k.a. *rational canonical normal form*) of M. Only the first of these – the computation of cyclic vectors – will be described in this summary. Details can be found in [1].

In the sequel, we will restrict to the case where the characteristic polynomial of A is squarefree, hence equal to the minimal polynomial of A.

2. Cyclic vectors and companion matrices

Let $P(X)$ be a monic polynomial of degree n with coefficients in k:

$$P(X) = X^n + \sum_{i=0}^{n-1} p_i X^i.$$

It is easy to see that $P(X)$ is the characteristic polynomial of the so-called *companion matrix*

$$C_P = \begin{pmatrix} 0 & 0 & \cdots & 0 & -p_0 \\ 1 & 0 & \cdots & 0 & -p_1 \\ 0 & 1 & \cdots & 0 & -p_2 \\ & \ddots & & & \\ 0 & & \cdots & 1 & -p_{n-1} \end{pmatrix}.$$

Let A be a linear operator over k and let $P_A(X)$ denote its minimal polynomial.

Definition 1. If v is a vector in k^n, the minimal polynomial of A relatively to v is the lowest degree nonzero polynomial $P_v(X)$ such that $P_v(A)v = 0$.

Definition 2. A vector v is called *cyclic* if and only if $P_v(X) = P_A(X)$.

One has the following:

Theorem 1. *Every linear operator A has a cyclic vector.*

In the case where P_A is equal to the characteristic polynomial, and if v is a cyclic vector, the matrix of A in the basis $(v, Av, \dots, A^{n-1}v)$ is a companion matrix.

Let M be the matrix of A and call C the companion matrix of its characteristic polynomial. The problem we want to solve is how to compute a cyclic vector as fast as possible. We will perform this operation in several steps: the first one is the computation of the Shift-Hessenberg form of M, noted H, and the second is finding C from H.

3. The Shift-Hessenberg form

Proposition 1. *Let M be an $n \times n$ matrix of $\mathcal{M}_n(k)$. There exists a matrix H similar to M of the form:*

$$\begin{pmatrix}
0 & 0 & \times & & & & \times & & & \times \\
1 & 0 & \times & & & & \times & & & \times \\
0 & 0 & \times & & & & \times & & & \times \\
0 & 1 & 0 & & & & \times & & & \times \\
0 & 0 & & 1 & & & \times & & & \times \\
 & & & & \ddots & & \cdots & & & \cdots \\
 & & & & & 1 & \times & & & \times \\
 & & & & & & 0 & & & \times \\
 & & & & & & & 1 & & \times \\
 & & & & & & & & \ddots & \cdots \\
 & & & & & & & & & \times
\end{pmatrix}.$$

The matrix H is called the Shift-Hessenberg form *of M (SHS for short). Computation of H requires $O(n^3)$ elementary operations in k.*

Proof. Do as in Gauss reduction, but starting from the sub-diagonal. If there is no non-zero element in the first column, below the sub-diagonal, then do nothing. Otherwise, assuming that it is $M_{1,2}$ (permuting lines if needed), eliminate all non-zero entries of this column. At the end of the process, we end up with a matrix of the above form.

The cost of this algorithm is very close to that of Gaussian reduction, that is $O(n^3)$. $\square$

It is clear, that when we can find a pivoting element for each column, we end up with a companion matrix. More generally, any SHS matrix can be written as

$$H = \begin{pmatrix}
H_{B_1,B_1} & H_{B_1,B_2} & \cdots & H_{B_1,B_m} \\
0 & H_{B_2,B_2} & \cdots & H_{B_2,B_m} \\
\vdots & & \ddots & \vdots \\
0 & \cdots & 0 & H_{B_m,B_m}
\end{pmatrix} \tag{1}$$

where m is an integer called the *parameter* of H and where each H_{B_i,B_j} is a companion matrix. It can be shown that the minimal polynomials of the H_{B_i,B_j} are pairwise coprime.

4. From the SHS form to the companion form

The key of the algorithm is the following Lemma [1]

LEMMA 1. *Let A be any block-triangular matrix with two blocks:*

$$A = \begin{pmatrix} A_{B_1,B_1} & A_{B_1,B_2} \\ 0 & A_{B_2,B_2} \end{pmatrix}.$$

For $i = 1, 2$, let $f_i(X)$ be the minimal polynomial of A_{B_i,B_i}. Assume f_1 and f_2 are relatively prime. Let v_{B_i} be a cyclic vector for A_{B_i,B_i}. Let h_2 be such that $h_2(X)f_2(X) \equiv 1 \bmod f_1(X)$. Then a cyclic vector for A is given by

$$v = \begin{pmatrix} u_{B_1} \\ v_{B_2} \end{pmatrix}$$

where

$$u_{B_1} = h_2(A_{B_1,B_1})\left((f_2(A)u_{B_2})_{|B_1} - v_{B_1}\right).$$

The computation of v can be done in $O(n^3)$ field operations.

Now suppose we are given H in the form of (1). If $m = 1$, H is already a companion matrix. If $m = 2$, the preceding Lemma applies. When $m > 2$, there exist two strategies. The first one is to compute a cyclic vector for the last two blocks, replacing these blocks by a companion matrix, and so on, until the whole matrix is companion. The second one is to split H in the form of

$$(2) \qquad \begin{pmatrix} H_{B_1,B_1} & H_{B_1,B_2} \\ 0 & H_{B_2,B_2} \end{pmatrix}$$

such that the sizes of H_{B_1,B_1} and H_{B_2,B_2} are kept under control. These can be chosen such that either H_{B_1,B_1} is a single companion block of size $\geq 2n/3$ or both matrices have size $\leq 2n/3$. This leads to two deterministic algorithms. The first one is iterative and has cost $O(n^3 + n^2m^2)$; the second one is recursive and has cost $O(n^3)$. We note that on average, the parameter m is $O(\log n)$.

All these algorithms have been implemented in AXIOM and give very encouraging running times.

5. Normal bases

DEFINITION 3. Let K be a finite extension of degree n of k. An element $\alpha \in K$ is *normal* if and only if

$$K = \mathrm{Vect}_k(\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{n-1}}).$$

If α is normal, then $(\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{n-1}})$ is called a *normal basis.*

Using a normal basis is particularly useful when computing powers of elements, since this is readily done via a cyclic shift:

$$(a_0, a_1, \ldots, a_{n-1})^q = (a_{n-1}, a_0, \ldots, a_{n-2}).$$

Moreover, it is easy to construct a multiplication table for k by precomputing the quantities $\alpha \times \alpha^{q^i}$ for all i. We can see this as follows. Write

$$c = \left(\sum_{i=0}^{n-1} a_i \alpha^{q^i}\right)\left(\sum_{i=0}^{n-1} b_i \alpha^{q^i}\right) = \left(\sum_{i=0}^{n-1} c_i \alpha^{q^i}\right).$$

Then $c_0 = F_0(a, b)$ is a bilinear symmetric form. Using $c^{q^{n-1}} = a^{q^{n-1}} b^{q^{n-1}}$ we deduce that

$$(a_1\alpha + a_2\alpha^q + \cdots)(b_1\alpha + b_2\alpha^q + \cdots) = c_1\alpha + \cdots$$

or $c_1 = F_0(a^\sigma, b^\sigma)$ where σ denotes the shift operation. We see that computing all c_i's needs only one matrix operation, followed by conjugation.

It is easy to see that:

PROPOSITION 2. *Let $\pi : x \mapsto x^q$ denote the Frobenius automorphism. Then α is normal if and only if α is a cyclic vector for π.*

Using the results of the preceding sections, and noting that the minimal polynomial of π is $X^n - 1$, that is squarefree for $(n, p) = 1$, we get

THEOREM 2. *We can find a normal element in deterministic time $O(n^3 + n^2 \log q)$, where the last term accounts for the computation of a matrix representing π.*

This result improves upon earlier results by von zur Gathen and Giesbrecht who gave a probabilistic algorithm in $\mathcal{O}(n^2 \log q)$ (using fast polynomial multiplication) or $O(n^3 \log q)$ without fast multiplication, and a deterministic algorithm running in time $O(n^4 + n^2 \log q)$.

It is possible to treat along the same lines the case where $n = p^k$.

Bibliography

[1] Augot (Daniel) and Camion (Paul). – *On the computation of minimal polynomial, cyclic vectors and the Frobenius form.* – Research Report n° 2006, INRIA, August 1993.

[2] Giesbrecht (Mark William). – *Nearly optimal algorithms for canonical matrix forms.* – PhD thesis, University of Toronto, 1993.

[3] Lidl (Rudolf) and Niederreiter (Harald). – *Finite Fields.* – Addison-Wesley, 1983, *Encyclopedia of Mathematics and its Applications*, vol. 20.

[4] McEliece (Robert). – *Finite fields for computer scientists and engineers.* – Kluwer Academic Publishers, 1988, *Kluwer international series in engineering and computer science.*

[5] Menezes (Alfred J.). – *Applications of Finite Fields.* – Kluwer Academic Publishers, 1993.

[6] Ozello (Patrick). – *Calcul exact des formes de Jordan et de Frobenius d'une matrice.* – Thèse, Université Scientifique Technologique et Médicale de Grenoble, 1987.

[7] von zur Gathen (J.) and Giesbrecht (M.). – Constructing normal bases in finite fields. *Journal of Symbolic Computation*, vol. 10, 1990, pp. 547–570.