W

Wiedemann

Let

- k be a field,
- $A \in K^{n\times n}$ an $n \times n$ square matrix over k,
- $b \in K^n$ any column vector and
- $u \in K^{1\times n}$ any row vector;

Denoting

- $f(\lambda) := 1 + \sum_{i=1}^{n} f_i\lambda^i \in k[\lambda]$ the *minimum polynomial* of A, normalized so that $f(0) = 0$

we remark that $A^\ell = -\sum_{i=1}^{n} f_i A_i^{i+\ell}$, for each $\ell \in \mathbb{N}$. As a consequence

$$A^\ell b = -\sum_{i=1}^{n} f_i A_i^{i+\ell} b \text{ and } uA^\ell b = -\sum_{i=1}^{n} f_i u A_i^{i+\ell} b, \text{ for each } \ell \in \mathbb{N}.$$

Remark W.1. In particular if, for each $i \in \mathbb{N}$ we denote

$$A^i := \left(a_{jh}^{(i)}\right), A^i b = \left(b_1^{(i)}, \ldots, (b_1^{(n)}\right)^T,$$

and

$$c_i := uA^i b \in K$$

both each sequences $a_{jh}^{(0)}, a_{jh}^{(1)}, \ldots, a_{jh}^{(i)}, \ldots, 1 \le j, h \le n$, $b_j^{(0)}, b_j^{(1)}, \ldots, b_j^{(i)}, \ldots, 1 \le j \le n$, and, mainly

$$c_0, c_1, \ldots, c_i, \ldots$$

are *linearly recurring sequences* owning a *minimal polynomial*. □

It is then worthwhile to extend such notion to vectorspaces:

Definition W.2. Let V be a vector space and let $a_0, \ldots, a_i, \ldots$ be an infinite sequence with elements $a_i \in V$. The sequence is *linearly generated over* K if there is polynomial $f(\lambda) := \sum_{i=0}^{n} f_i\lambda^i \in k[\lambda] \setminus \{0\}$ s.t. $\sum_{i=0}^{n} f_i a_{=}0$. Any such polynomial is called a *generating polynomial* for the sequence.

It is then clear that the set of all generating polynomials for the sequence, together with the zero polynomial forms an ideal.

Definition W.3. The unique polynomial generating such ideal is called the *minimum polynomial* of the sequence.

On the basis of this definition we will denote

- $f^{A,b}(\lambda) \in k[\lambda]$ the minimum polynomial of the sequence $b, Ab, A^2b, \ldots, A^ib, \ldots$, for each column vector $b \in K^n$ and
- $f_u^{A,b}(\lambda) \in k[\lambda]$ the minimum polynomial of the sequence $c_0 := u \cdot b, c_1 := uAb, c_2 := uA^2b, \ldots, c_i := uA^ib, \ldots$, for each column vector $b \in K^n$ and each row vector $u \in K^{1\times n}$.

remarking that $f_u^{A,b} \mid f^{A,b} \mid f$.

Remark W.4. Usually the minimal polynomial is normalized in order to have leading coefficient 1.

Inthe context of Wiedermann Algorithm (but also of Berlekamp–Massey Algorithm) it is better to normalize it in order to have trailing coefficient 1. □

W.1 The nonsingular case

Let us now assume A to be nonsingular, so that for each $b \in K^n \setminus \{\mathbf{0}\}$, there is a unique $x \in K^n \setminus \{\mathbf{0}\}$, such that

$$Ax = b.$$

In this case, denoting

$$f^{A,b}(\lambda) := 1 + \sum_{i=1}^{n} f_i \lambda^i \in k[\lambda]$$

the minimum polynomial of the sequence $b, Ab, A^2b, \ldots, A^ib, \ldots$ we have

$$f^{A,b}(Ab) := b + \sum_{i=1}^{n} f_i A^i b = 0 \implies b = -\sum_{i=1}^{n} f_i A^i b = A\left(-\sum_{i=1}^{n} f_i A^{i-1} b\right)$$

so that $x := -\sum_{i=1}^{n} f_i A^{i-1} b$ is the required solution.

W.2 The singular case

Let us now assume A to be singular. In this case the minimum polynomial $f^{A,b}(\lambda)$ of the sequence $b, Ab, A^2b, \ldots, A^ib, \ldots$ is such that $f^{A,b}(0) = 0$ and let

$$f^- := f^{A,b}(\lambda)/\lambda = \sum_{i=1}^{n} f_i \lambda^{i-1} \in k[\lambda].$$

Then $x := f^-(A)b$ satisfies

$$A \cdot x = A \cdot \left(\sum_{i=1}^{n} f_i A^{i-1} b\right) = \sum_{i=1}^{n} f_i A^i b = f^{A,b}(A) \cdot b = 0$$

W.3 Computing the minimal polynomnial

In order to compute the minimal polynomial, one can apply the Berlekamp–Massey Algorithm to the sequence

$$c_0 := u \cdot b, c_1 := uAb, c_2 := uA^2b, \ldots, c_i := uA^ib, \ldots$$

where $u \in K^{1\times n}$ is any random row vector.

The algorithm returns the minimal polynomial $f_u^{A,b} \mid f^{A,b}$ If $b_1 := f_u^{A,b}(Ab) = f_u^{A,b}(A) \cdot b = 0$ then $f_u^{A,b} = f^{A,b}$ and we are throu.

If instead, $b_1 \neq 0$ we have in any case found a factor of $f^{A,b}$. In this case one can reapply the same procedure with a different random row vector u', but it is more efficient to apply the Berlekamp–Massey Algorithm to the sequence

$$c_0 := u_1 \cdot b_1, c_1 := u_1 Ab_1, c_2 := u_2 A^2 b_2, \ldots, c_i := u_2 A^i b_2, \ldots$$

where $u_1 \in K^{1\times n}$ is any random row vector, obtaining the minimal polynomial $f_{u_1}^{A,b_1}$. If

$$b_2 := f_{u_1}^{A,b_1}(Ab_1) = f_{u_1}^{A,b_1}(A) \cdot b_1 = f_{u_1}^{A,b_1}(A) \cdot f_u^{A,b}(A) \cdot b = \left(f_{u_1}^{A,b_1} f_u^{A,b}\right)(A) \cdot b = 0$$

then $f^{A,b}(\lambda) = f_{u_1}^{A,b_1}(\lambda) f_u^{A,b}(\lambda)$; otherwise we repeat the same procedure with b_2 and a new random row vector $u_2 \in K^{1\times n}$.

Eventually we will obtain the case in which

$$\begin{aligned} b_{k+1} &:= f_{u_k}^{A,b_k}(Ab_k) \\ &= f_{u_k}^{A,b_k}(A) \cdot b_k \\ &= f_{u_k}^{A,b_k}(A) f_{u_{k-1}}^{A,b_{k-1}}(A) \cdots f_u^{A,b}(A) \cdot b \\ &= \left(f_{u_k}^{A,b_k} f_{u_{k-1}}^{A,b_{k-1}} \cdots f_{u_1}^{A,b_1} f_u^{A,b}\right)(A) \cdot b \\ &= 0 \end{aligned}$$

so that $f^{A,b}(\lambda) = f_{u_k}^{A,b_k}(\lambda) f_{u_{k-1}}^{A,b_{k-1}}(\lambda) \cdots f_{u_1}^{A,b_1}(\lambda) f_u^{A,b}(\lambda)$.

References

[W] D.H.Wiedemann, *Solving Sparse Linear Equations over Finite Fields*, IEEE Trans. on Inf. Th. **32** (1986), 54–62.