Proof for the Smith form

Lemma 1 (Polynomial division)**.** *Let* $a, b \in \mathbb{C}[\lambda]$ *be two polynomials with* $b \neq 0$ *and* $\deg a \geq \deg b \geq 0$. *Then there exist unique* $q, r \in \mathbb{C}[\lambda]$ *such that*

$$a = qb + r \tag{1}$$

with $\deg r < \deg b$. *If* $\deg a \geq \deg b$ *we also have* $\deg q = \deg a - \deg b$.

Proof. The result follows from division with remainder. If $\deg a < \deg b$ then the statement is trivial. For the other case, we give an inductive proof.

Base case: $\deg a = \deg b$ in this case a and b have the form $a(\lambda) = \sum_{i=0}^{K} a_i \lambda^i$ and $b(\lambda) = \sum_{i=0}^{K} b_i \lambda^i$, with $a_K, b_K \neq 0$. Define $q \in \mathbb{C}[\lambda]$ as a constant $q(\lambda) := \frac{a_K}{b_K}$. Then

$$(qb)(\lambda) = \sum_{i=0}^{K} \left(b_i \frac{a_K}{b_K} \right) \lambda^i$$

and thus $r \in \mathbb{C}[\lambda]$ defined by

$$r(\lambda) := (a - qb)(\lambda) = \sum_{i=0}^{K} \left(a_i - b_i \frac{a_K}{b_K} \right) \lambda^i = \sum_{i=0}^{K-1} \left(a_i - b_i \frac{a_K}{b_K} \right) \lambda^i$$

is a polynomial of degree $\leq K - 1 < K$. Also we have $\deg q = 0 = K - K = \deg a - \deg b$.

Inductive step: Write a and b in the form $a(\lambda) = \sum_{i=0}^{K+M} a_i \lambda^i$ and $b(\lambda) = \sum_{i=0}^{K} b_i \lambda^i$, with $M \in \mathbb{N}$ and $a_{K+M}, b_K \neq 0$. Define $q_0 \in \mathbb{C}[\lambda]$ by $q_0(\lambda) := \lambda^M \frac{a_{K+M}}{b_K}$. Then

$$(q_0 b)(\lambda) = \sum_{i=0}^{K} b_i \frac{a_{K+M}}{b_K} \lambda^{i+M} = \sum_{i=M}^{K+M} b_{i-M} \frac{a_{K+M}}{b_K} \lambda^i$$

and thus $r_0 \in \mathbb{C}[\lambda]$ defined by

$$r_0(\lambda) := (a - q_0 b)(\lambda) = \sum_{i=M}^{K+M} \left(a_i - b_{i-M} \frac{a_{K+M}}{b_K} \right) \lambda^i + \sum_{i=0}^{M-1} a_i \lambda^i = \sum_{i=M}^{K+M-1} \ldots + \sum_{i=0}^{M-1} \ldots,$$

is a polynomial of degree $\leq K + M - 1$. Using the induction hypothesis we conclude the existence of $q_1, r \in \mathbb{C}[\lambda]$ which fulfill $r_0 = q_1 b + r$, $\deg r < \deg b$, and $\deg q_1 = \deg r_0 - \deg b = (K + M - 1) - K = M - 1$. Setting $q := q_0 + q_1$ we find that

$$a = q_0 b + r_0 = q_0 b + q_1 b + r = (q_0 + q_1) b + r = qb + r.$$

Since $\deg q_1 = M - 1 \leq \deg q_0$ we also have $\deg q = \deg q_0 = M = (K + M) - K = \deg a - \deg b$.

For uniqueness, let $q, r \in \mathbb{C}[\lambda]$ and $\tilde{q}, \tilde{r} \in \mathbb{C}[\lambda]$ both fulfill (1). Then we have

$$(r - \tilde{r}) + b(q - \tilde{q}) = 0. \tag{2}$$

If $q - p$ was nonzero, then $\deg b(q - \tilde{q}) \geq \deg b > \deg r - \tilde{r}$ which contradicts (2). This implies $q = \tilde{q}$ which again by (2) implies $r = \tilde{r}$. □

Definition 2. We say that $b \in \mathbb{C}[\lambda]$, $b \neq 0$ divides $a \in \mathbb{C}[\lambda]$ if in Lemma 1 we have $r = 0$.

Theorem 3 (Smith canonical form)**.** *Let $P \in \mathbb{C}[\lambda]^{p,q}$. Then there exists an $r \in \mathbb{N}_0$ and unimodular matrices $S \in \mathbb{C}[\lambda]^{p,p}$, $T \in \mathbb{C}[\lambda]^{q,q}$ such that*

$$P = S \begin{bmatrix} \operatorname{diag}(d_1, \ldots, d_r) & 0 \\ 0 & 0 \end{bmatrix} T$$

where $d_1, \ldots, d_r \in \mathbb{C}[\lambda]$ are polynomials with $d_i \neq 0$ for $i = 1, \ldots, r$ and d_{i+1} divides d_i for $i = 1, \ldots, r-1$.

Proof. (From [PW98, Theorem B.1.4]) The proof is an algorithm. Assume that P is nonzero, since otherwise the statement is trivial. By $\operatorname{mindeg}(P)$ we denote the minimal degree of all the nonzero elements of P. In the following we are going to apply a series of unimodular pre- and post-multiplications to P until P has Smith form. Since the product of unimodular matrices is again unimodular we have then indeed found the Smith form. To simplify notation we will in the following write P, whenever actually the matrix is meant which arises from P by the proclaimed unimodular pre- and post-multiplications. With this convention, the algorithm is finished if P is in Smith form. A will denote the elements of P by $p_{i,j}$.

a) Apply row and column permutations to P to achieve that a nonzero element with degree $\operatorname{mindeg}(P)$ appears at the $(1,1)$ position. Using Lemma 1 (with $b = p_{1,1}$ and $a = p_{i,1}$), we obtain $q_{i,1}$ and $r_{i,1}$ such that $p_{i,1} = q_{i,1}p_{1,1} + r_{i,1}$ for $i = 2, \ldots, p$. Then we have

$$\underbrace{\begin{bmatrix} 1 & & & \\ -q_{2,1} & 1 & & \\ \vdots & & \ddots & \\ -q_{p,1} & & & 1 \end{bmatrix}}_{=:Q_0} \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,q} \\ p_{2,1} & p_{2,2} & \cdots & p_{1,q} \\ \vdots & \vdots & & \vdots \\ p_{p,1} & p_{p,2} & \cdots & p_{p,q} \end{bmatrix} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,q} \\ r_{2,1} & \star & \cdots & \star \\ \vdots & \vdots & & \vdots \\ r_{p,1} & \star & \cdots & \star \end{bmatrix}$$

where the $\star$-entries denote polynomials which are not further specified and Q_0 is unimodular. Similar, by a post-multiplication with a unimodular matrix one can achieve that all entries in the $(1,j)$ positions (with $j = 2, \ldots, q$) have degree smaller than $p_{1,1}$. If we do not have

$$p_{i,1} = 0 \text{ for } i = 2, \ldots, p \text{ and } p_{1,j} = 0 \text{ for } j = 2, \ldots, q, \tag{3}$$

then $\operatorname{mindeg}(P)$ has at least decreased. In this case goto a). If $\operatorname{mindeg}(P) = 0$ at the beginning of a), then at the end, condition (3) will be fulfilled in any case. Thus, since degrees (of nonzero polynomials) are nonnegative and $\operatorname{mindeg}(P)$ decreases in each repetition of a), we see that (3) is fulfilled after a finite number of steps; then goto b).

b) We have reached the situation (3). Either $p_{1,1}$ divides all the other elements of P, or there exists a column that contains an element that is *not* a multiple of the $(1,1)$ element. If the latter is true, add this column to the first column of P and start again at a). Because of Definition 2 this will decrease $\operatorname{mindeg}(P)$ in the first step. Thus, after a finite number of repetitions of b), we have that the element $p_{1,1}$ divides all other elements, since this holds at the latest when $\operatorname{mindeg}(P) = 0$.

c) We have reached the situation (3) such the the element $p_{1,1}$ divides all other entries of P. Factor out the common divisor from P (call it d_1) such that the $(1,1)$ element of P becomes a nonzero constant. Then, in an inductive fashion, start again at a) with the matrix obtained from P by deleting the first row and column.

□

One can show that the quantities $d_1, \ldots, d_r$ in the Smith form are unique, see [Gan59, p. 139, §3].

References

[Gan59] F.R. Gantmacher. *The Theory of Matrices I.* Chelsea Publishing Company, New York, NY, 1959.

[PW98] J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioral Approach.* Springer, Berlin, 1998.