A Matrix Euclidean Algorithm induced by State Space Realization*

Brian M. Allen
Department of Mathematics
University of Notre Dame
Notre Dame, Indiana 46556-5683, USA
e-mail: ballen1@nd.edu

Joachim Rosenthal
Department of Mathematics
University of Notre Dame
Notre Dame, Indiana 46556-5683, USA
e-mail: Rosenthal.1@nd.edu

July, 1998

Abstract

An efficient algorithm is developed for determining the greatest common left divisor (GCLD) of two polynomial matrices. Knowing this divisor allows for several immediate applications: In coding theory, a noncatastrophic convolutional encoder can be derived from an arbitrary one. In systems theory, irreducible matrix fraction descriptions of transfer function matrices can be found. In linear algebra, the greatest common divisor can be seen as a basis for a free module generated by the columns of the matrices.

The approach taken is based on recent ideas from systems theory. A minimal state space realization is obtained with minimal calculations, and from this the controllability matrix is analyzed to produce the GCLD. It will be shown that the derived algorithm is a natural extension of the Euclidean algorithm to the matrix case.

1 Introduction

Let $\mathbb{F}$ be an arbitrary field and consider the polynomial ring $\mathbb{F}[s]$. If we are given two polynomial matrices $E(s)$ and $F(s)$ each with p rows then we may define a greatest common left divisor (GCLD) to be any $p \times p$ polynomial matrix $L(s)$ satisfying:

1. There exists polynomial matrices $\tilde{E}(s)$ and $\tilde{F}(s)$ such that $L(s)\tilde{E}(s) = E(s)$ and $L(s)\tilde{F}(s) = F(s)$.

2. If $\bar{L}(s)$ is any other divisor of $E(s)$ and $F(s)$ then there exists a polynomial matrix $D(s)$ such that $\bar{L}(s)D(s) = L(s)$.

By an arbitrary choice, we will work with left divisors. The theory holds *mutatis mutandis* for right divisors.

*Both authors were supported in part by NSF grant DMS-96-10389.

Notice that GCLD's are not unique. For our applications we will assume that the matrix $[E(s) \mid F(s)]$ is full rank. This implies that all GCLD's will be nonsingular and differ by a unimodular right factor [11]. Note also that the columns of the GCLD of the full rank polynomial matrix $[E(s) \mid F(s)]$ form a basis for the free module spanned by the columns of $[E(s) \mid F(s)]$ in $\mathbb{F}^p[s]$. Two matrices are said to be coprime if their GCLD is a unimodular matrix.

Instead of starting with two separate matrices and then combining them into one, we are given a single full rank matrix $P(s)$ of size $p \times (p+m)$. We can speak of the GCLD of this single matrix by writing $P(s) = [E(s)|F(s)]$ where usually $E(s)$ is of size $p \times p$ and $F(s)$ is of size $p \times m$, and hence the GCLD of $P(s)$ is then the GCLD of $E(s)$ and $F(s)$. Obviously the GCLD does not depend on how we choose the division. Equivalently, we could define a GCLD of $P(s)$ to be a matrix $L(s)$ such that $L(s)\tilde{P}(s) = P(s)$, where $\tilde{P}(s)$ is a polynomial matrix whose Smith or equivalently Hermite form is $[I_p \mid 0]$.

With this last description we are able to see several immediate applications. First, if we are given $P(s)$ as a polynomial basis for a rational vector space [8], then by dividing by $L(s)$ (*i.e.* taking $\tilde{P}(s)$) we get a minimal polynomial basis for the vector space (as defined in [8]). Secondly, if we are given $P(s)$ as a generating set for its column module over $\mathbb{F}[s]$, then we observed earlier that the columns of the GCLD, $L(s)$, of $P(s)$ form a basis of the column module of $P(s)$. In particular if $P(s)$ is of size 1×2 and has the form $P(s) = (p(s), q(s)),\ p(s), q(s) \in \mathbb{F}[s]$ then the GCLD of $P(s)$ is nothing else than the greatest common divisor (g.c.d.) of $p(s), q(s)$. Moreover our algorithm is in this case equivalent to Euclid's algorithm. Finally, if we are given $P(s)$ as a convolutional encoder, then $P(s)$ is an observable (*i.e.* non-catastrophic with delay 0) encoder if and only if $L(s)$ is a unimodular matrix [1, 7].

Closely related to this last application, we can think of $P(s)$ as describing over the real numbers $\mathbb{R}$ a linear behavior in the sense of Willems [17]:

$$\mathfrak{B} = \{w(t) \in C^\infty(\mathbb{R}, \mathbb{R}^{m+p}) \mid P\left(\frac{d}{dt}\right) w(t) = 0\}.$$

The computation of the GCLD is then needed for the computation of the controllable subbehavior of $\mathfrak{B}$.

The approach that will be taken in this paper is to obtain a minimal state space representation of the associated behavior $\mathfrak{B}$ with little or no calculation [15]. This state space representation will be controllable if and only if our behavioral system (or encoder) is observable. Further, the contribution of this paper will be to calculate a GCLD of $P(s)$ directly from the controllability matrix of this state space representation.

As we shall see, the algorithm presented will be a natural generalization of the Euclidean algorithm to polynomial matrices. The algorithm has been induced on the state space level by an efficient Gaussian elimination and this explains our choice of title.

2 A Brief History of the Problem

The problem of finding GCLD's is not new, and, indeed, there are several algorithms in existence. The most obvious way is to append the two matrices together as $[E(s) \mid F(s)]$

and perform polynomial column operations (over $\mathbb{F}[s]$) to bring the matrix to Smith or Hermite form [4]. The obvious drawback is that polynomial column operations can become quite tedious, especially if the degrees of the polynomial entries are high. This problem was overcome by Kung *et al.* [12, 6] with their approach using generalized Sylvester matrices. A problem with that algorithm is that the scalar matrices obtained from the original polynomial matrices were often quite large.

Several more recent works, using somewhat similar but distinct methods to the one proposed here, have appeared: Fuhrmann [9] obtained an algorithm using a matrix continued fraction representation. Antoulas [3] has done considerable work on the subject using recursive and partial realizations.

An excellent reference on the various techniques of computing GCD's in the case $p = 1$ can be found in [5]. In fact, the section on "G.C.D. Using Companion Matrix" from this book give exactly our algorithm in the simple case $p = 1$. In this reference it was, unfortunately, not observed by the author that the companion matrix was, in fact, a realization of the polynomial matrix. This prevented the extension to the general case, where the author of that paper instead presents the algorithm of Kung *et al.*.

3 Realization

We now present the main result we will need, proceeded by some notation. For a more thorough account of the ideas involved, please refer to [15].

Partition $P(s)$ into $P(s) = [\, E(s) \mid F(s) \,]$, where $E(s)$ is $p \times p$ and $F(s)$ is $p \times m$. After some unimodular row operations we can assume that $P(s)$ is row proper with row degrees (Kronecker indices) $\nu_1 \geq \cdots \geq \nu_p$. After a possible right multiplication by a $(m+p)\times(m+p)$ invertible matrix we can assume that the high order coefficient matrix, P_h, has the form $[I_p \mid 0]$. Assume that $P(s)$ has no constant rows, i.e. $\nu_p \geq 1$. For $i, j = 1, \ldots, p$ let

$$e_{i,j}(s) \;=\; \sum_{\alpha=0}^{\nu_i} e_{i,j}^{\alpha} s^{\alpha} \qquad\qquad \mathbf{f}_i(s) \;=\; \sum_{\alpha=0}^{\nu_i - 1} \mathbf{f}_i^{\alpha} s^{\alpha}$$

denote the polynomial entries of E(s) and the i^{th} row of $F(s)$ respectively.

Define for $i = 1, \ldots, p$ matrices of sizes $\nu_i \times \nu_i$, $\nu_i \times m$ and $1 \times \nu_i$ respectively:

$$A_{i,i} := \begin{bmatrix} 0 & \ldots & \ldots & \ldots & -e_{i,i}^{0} \\ 1 & 0 & & & -e_{i,i}^{1} \\ 0 & 1 & \ddots & & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -e_{i,i}^{\nu_i - 1} \end{bmatrix}, \; B_i := \begin{bmatrix} \mathbf{f}_i^{0} \\ \mathbf{f}_i^{1} \\ \vdots \\ \mathbf{f}_i^{\nu_i - 1} \end{bmatrix}, \; C_i \;:=\; [0, \ldots, -1]. \tag{3.1}$$

For $i, j = 1, \ldots, p$, $i \neq j$ define matrices of size $\nu_i \times \nu_j$:

$$A_{i,j} \;:=\; \begin{bmatrix} 0 & \ldots & 0 & -e_{i,j}^{0} \\ \vdots & & \vdots & -e_{i,j}^{1} \\ \vdots & & \vdots & \vdots \\ 0 & \ldots & 0 & -e_{i,j}^{\nu_i - 1} \end{bmatrix} \tag{3.2}$$

The matrices $A_{i,i}$ are just the companion matrices for the polynomials $e_{i,i}(s)$, while the matrices $A_{i,j}$ are just $\nu_j - 1$ columns of zeroes with the coefficient vector of the polynomial $e_{i,j}(s)$ appended on the right. Similarly each B_i is just the coefficient vectors of all the polynomials in the i^{th} row of $F(s)$. So these matrices are obtained with no calculations at all, provided that the matrix $P(s)$ meets the somewhat stringent conditions imposed. If P_h does not have the form $P_h = [I_p \mid 0]$ then it can be brought into this form with the unimodular operations outlined above.

Notice also the requirement that $P(s)$ has no constant rows. If $P(s)$ has κ constant rows then the row and column operations outlined above will transform $P(s)$ into:

$$\hat{P}(s) = \left[\begin{array}{cc|c} \hat{E}_1(s) & \hat{E}_2(s) & \hat{F}(s) \\ 0 & I_\kappa & 0 \end{array}\right] \tag{3.3}$$

and $[\,\hat{E}_1(s) \mid \hat{F}(s)\,]$ has no constant rows.

Right unimodular operations will not affect the GCLD, however left operations will have to be 'undone' once the GCLD of the resulting matrix is calculated. So all of these conditions can be met at the expense of some efficiency. From here on, assume that $P(s)$ meets these requirements.

Theorem 3.1 ([18, 15]) *Given $P(s) = [\,E(s) \mid F(s)\,]$, satisfying $P_h = [\,I_p \mid 0\,]$ and let $A_{i,j}, B_i, C_i$ be defined as above. Let*

$$A := \begin{bmatrix} A_{1,1} & \cdots & A_{1,p} \\ \vdots & \ddots & \vdots \\ A_{p,1} & \cdots & A_{p,p} \end{bmatrix},\ B := \begin{bmatrix} B_1 \\ \vdots \\ B_p \end{bmatrix},\ C := \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_p \end{bmatrix}$$

and let σ represent either the shift operator or the differential operator $\frac{d}{dt}$. Then

$$\begin{array}{rcl} \sigma x(t) & = & Ax(t) + Bu(t), \\ y(t) & = & Cx(t) \end{array} \tag{3.4}$$

represents a minimal state space realization of the system

$$E(\sigma)y(t) + F(\sigma)u(t) = 0. \tag{3.5}$$

In particular (3.5) *represents a controllable behavior if and only if (A, B) is a controllable pair, i.e. the controllability matrix*

$$\mathcal{C}(A, B) := \begin{bmatrix} B & AB & \ldots A^{n-1}B \end{bmatrix}$$

has full rank.

As usual, we call (3.4) an (A, B, C) representation of the system (3.5). We see that A has size $n \times n$ (where $n = \sum_{i=1}^{p} \nu_i$), B is $n \times m$, and C is $p \times n$.

The idea here is that controllability of the state space representation is equivalent to the controllability of the behavioral system given by $P(s)$ which is equivalent to $P(s)$ being an observable encoder [14, 15].

The relationship between the polynomial matrix $P(s)$ and the matrices (A, B, C) is expressed in the following way: Consider the $p \times n$ matrix

$$X(s) = \begin{bmatrix} 1 & s & \cdots & s^{\nu_1-1} & 0 & 0 & \cdots & 0 & & & & & \\ 0 & 0 & \cdots & 0 & 1 & s & \cdots & s^{\nu_2-1} & & & & & \\ & & & & 0 & 0 & \cdots & 0 & & & & & \\ & & & & & & & & \ddots & 0 & 0 & \cdots & 0 \\ & & & & & & & & & 1 & s & \cdots & s^{\nu_p-1} \end{bmatrix} \tag{3.6}$$

which was called a basis matrix of size $\nu = [\nu_1, \ldots, \nu_p]$ in [15] since it has the property that every polynomial p-vector $\varphi(s) \in \mathbb{F}^p[s]$ whose i-th component has degree at most $\nu_i - 1$ can uniquely be described through $\varphi(s) = X(s)\alpha, \alpha \in \mathbb{F}^n$.

A direct calculation reveals that $P(s)$ and the matrices (A, B, C) are related by:

$$X(s)[sI - A \mid B] = P(s) \begin{bmatrix} -C & 0 \\ O & I_m \end{bmatrix} \tag{3.7}$$

Of course, we can multiply $X(s)$ by an invertible matrix $S \epsilon GL_n$, on the right and obtain the equivalent realization $(S^{-1}AS, S^{-1}B, CS)$. We will make use of this fact in Section 5 to obtain a more suitable realization.

4 The Controllability Space

We are given a $p \times (p + m)$ full rank polynomial matrix $P(s)$ and wish to determine its GCLD, $L(s)$. Write $P(s) = L(s)\tilde{P}(s)$, where $\tilde{P}(s)$ has Smith form $[I_p \mid 0]$. We will assume that the rows of $\tilde{P}(s)$ form a minimal basis in the sense of Forney [8]. The row degrees $(\mu_1, \ldots, \mu_p)$ of $\tilde{P}(s)$ are therefore the minimal indices of the rational vector space generated by the rows of the matrix $P(s)$. We will not assume that $(\mu_1, \ldots, \mu_p)$ are ordered by size. Also write $P(s) = [E(s) \mid F(s)]$ and let $P_h = [I_p \mid 0]$ be the high order coefficient matrix.

Since $L(s)$ is determined uniquely up to unimodular right multiplication, we have a choice as to which $L(s)$ to work with, and hence which $\tilde{P}(s)$ to work with. The following lemma relates $L(s)$ and $\tilde{P}(s)$ and it singles out a nice choice:

Lemma 4.1 *If the rows of $\tilde{P}(s)$ form a minimal basis having row degrees $\mu_1, \ldots, \mu_p$ then $L(s)$ is uniquely determined from the identity $P(s) = L(s)\tilde{P}(s)$. The (i, j)-entry of $L(s)$ has degree at most $(\nu_i - \mu_j)$ or the entry is zero.*

It is possible to choose $\tilde{P}(s)$ such that the scalar matrix L_∞ whose (i, j)-entry is the coefficient of $s^{\nu_i - \mu_j}$ in the (i, j)-entry of $L(s)$ is lower triangular.

Proof: The first part of the lemma is a direct consequence of [8]. The second part will be established by induction. Using elementary column operations on $L(s)$ (this corresponds to elementary row operations on $\tilde{P}(s)$) it will be possible to eliminate all entries of the first row of L_∞ with the exception of one entry. After a possible permutation of the columns we can assume that the first row of L_∞ has with the exception of the entry $(1, 1)$ all entries equal to zero. Proceeding inductively row by row will establish the claim. □

Let $\tilde{P}_h$ be the high order coefficient matrix of $\tilde{P}(s)$. From the fact that both $\tilde{P}_h$ and P_h have rank p and from the identity $P_h = L_\infty \tilde{P}_h$ it follows that L_∞ is invertible. As a direct consequence we have:

Lemma 4.2 *Let $d := \sum \mu_i$ be the McMillan degree of $\tilde{P}(s)$. Then*

$$\deg \det L(s) = n - d = n - \operatorname{rank} \mathcal{C}(A, B).$$

Lemma 4.2 establishes a first relation between the GCLD, $L(s)$, and the controllability matrix $\mathcal{C}(A, B)$. It should be noted that this result, for the case $p = 1$, was known already in 1950 by MacDuffee [13] if not earlier. It is the goal of this and the next section to show that under certain conditions it is possible to compute $L(s)$ from the column space of $\mathcal{C}(A, B)$, i.e. from the reachability space of (A, B).

Since the high order coefficient matrix of $P(s)$ has the form $[I_p \mid 0]$ we can realize $P(s)$ by inspection to obtain the scalar matrices A, B, and C relative to the basis matrix $X(s)$. Hence the following equation holds:

$$X(s)[sI - A \mid B] \;=\; [E(s)C \mid F(s)] \tag{4.1}$$

Note that in order to realize $\tilde{P}(s)$, we need $\tilde{P}_h = [I_p \mid 0]$ and that the row degrees, μ_i, of $\tilde{P}(s)$ are at least one. To satisfy the first requirement, in general, we will have to multiply $\tilde{E}(s)$ by T to obtain a realizable form. *i.e.* $\tilde{P}_r(s) = [\tilde{E}(s)T \mid \tilde{F}(s)]$. The second requirement cannot be guaranteed. For this section and the following one we will assume that $\tilde{P}(s)$ has no constant rows, and the case where there are constant rows is considered in Section 6.

Now, realize $\tilde{P}_r(s)$ to obtain matrices $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$, relative to the canonical basis matrix $\tilde{X}(s)$. Hence the following equation holds:

$$\tilde{X}(s)[sI - \tilde{A} \mid \tilde{B}] \;=\; [\tilde{E}(s)T\tilde{C} \mid \tilde{F}(s)] \tag{4.2}$$

The controllability matrix of the pair $(\tilde{A}, \tilde{B})$ can also be computed. However, the usual definition of the controllability matrix is that of a $d \times dm$ matrix, where $\tilde{B}$ is of size $d \times m$. We can, however, naturally extend the size of this matrix to $d \times nm$. This is necessary for the following key result.

Theorem 4.3

$$L(s)\tilde{X}(s)\mathcal{C}(\tilde{A}, \tilde{B}) \;=\; X(s)\mathcal{C}(A, B)$$

Proof: Repeated applications of (4.1) give:

$$X(s)\mathcal{C}(A, B) = \left[F \mid sF - ECB \mid \cdots \mid s^{n-1}F - s^{n-2}ECB - s^{n-3}ECAB - \cdots - ECA^{n-2}B\right] \tag{4.3}$$

Repeated applications of (4.2) give:

$$L(s)\tilde{X}(s)\mathcal{C}(\tilde{A}, \tilde{B}) = \left[F \mid sF - ET\tilde{C}\tilde{B} \mid \cdots \mid s^{n-1}F - s^{n-2}ET\tilde{C}\tilde{B} - s^{n-3}ET\tilde{C}\tilde{A}\tilde{B} - \cdots - ET\tilde{C}\tilde{A}^{n-2}\tilde{B}\right].$$

By examining the above expressions, it is clear that the only step remaining in the proof is to show that $CA^iB = T\tilde{C}\tilde{A}^i\tilde{B}$ for all non-negative integer i.

We notice that $(sI-A)^{-1} = \sum_{i=0}^{\infty} \frac{A^i}{s^{i+1}}$. Starting with the equation $X(s)(sI-A) = E(s)C$, we apply this inverse to obtain $X(s) = E(s)\sum_{i=o}^{\infty} \frac{CA^i}{s^{i+1}}$ Further:

$$F(s) = X(s)B = E(s)\sum_{i=o}^{\infty} \frac{CA^iB}{s^{i+1}}.$$

Similarly, we have the equations:

$$\tilde{F}(s) = \tilde{X}(s)\tilde{B} = \tilde{E}(s)T\sum_{i=o}^{\infty} \frac{\tilde{C}\tilde{A}^i\tilde{B}}{s^{i+1}}.$$

Multiplying the last equation by $L(s)$ results in

$$F(s) = E(s)\sum_{i=o}^{\infty} \frac{T\tilde{C}\tilde{A}^i\tilde{B}}{s^{i+1}}.$$

Since $E(s)$ has high order coefficient matrix I_p, the columns of $E(s)$ are linearly independent over $\mathbb{F}[s]$ and we get:

$$\sum_{i=o}^{\infty} \frac{CA^iB}{s^{i+1}} \;=\; \sum_{i=o}^{\infty} \frac{T\tilde{C}\tilde{A}^i\tilde{B}}{s^{i+1}}$$

Equating coefficients in the above expression gives us the desired equality and completes the proof. □

This theorem is the key to the entire algorithm as the following corollary shows.

Corollary 4.4 *There exists an invertible matrix $W \in Gl_{mn}$ such that*

$$X(s)\mathcal{C}(A,B)W = \left[s^{\mu_1-1}\mathbf{l}_1 \ldots s\mathbf{l}_1\ \mathbf{l}_1 \mid \ldots \mid s^{\mu_p-1}\mathbf{l}_p \ldots s\mathbf{l}_p\ \mathbf{l}_p \;\mid\; O_{p\times(mn-d)}\right]. \tag{4.4}$$

In this representation the $p \times p$ matrix

$$L(s) = [\mathbf{l}_1, \ldots, \mathbf{l}_p]$$

represents a greatest common left divisor of $[E(s)\ F(s)]$ and $\mu_1, \ldots, \mu_p$ are the row degrees of $[\tilde{E}(s)\ \tilde{F}(s)]$.

Proof: Since $\tilde{P}_r(s)$ is a minimal basis, its realization, $(\tilde{A}, \tilde{B})$, must be a controllable pair. Therefore, there exists a scalar matrix $W \in Gl_{mn}$ such that $\mathcal{C}(\tilde{A}, \tilde{B})W = \left[I_d \mid O_{p\times(mn-d)}\right]$. Hence, the theorem implies that $X(s)\mathcal{C}(A,B)$ is column equivalent to a matrix whose columns are exactly the columns of a GCLD and also multiples of these columns (as the multiplication $\tilde{X}(s)\mathcal{C}(\tilde{A}, \tilde{B})$ indicates). □

5 The Refining Algorithm

By Theorem 4.3 and Corollary 4.4, the columns of $L(s)$ are contained in a matrix that is column equivalent (over $\mathbb{F}$) to $X(s)\mathcal{C}(A,B)$. The question is now, how to select these p columns of $L(s)$ from the nm columns of the controllability matrix? The answer is fairly simple: column reduce and then choose the appropriate p columns in a manner that will be described below. However, we must first reconsider our choice of basis matrix $X(s)$. The reason we have started with the one we have chosen is that it allows us to write down the matrices A and B a little easier. The downside is that when we column reduce the controllability matrix we start by eliminating the lower degree terms of the polynomials in row 1 of the corresponding matrix $X(s)\mathcal{C}(A,B)$. It would make much more sense to start eliminating the highest degree terms in each row. We accomplish this by replacing the standard basis matrix $X(s)$ introduced in (3.6) with the basis matrix

$$\mathfrak{X}(s) \;=\; \begin{bmatrix} s^{\nu_1-1} & 0 & & & s^{\nu_1-2} & 0 & & & \cdots & s^{\nu_1-\nu_1} & 0 & & 0 \\ 0 & s^{\nu_2-1} & & & 0 & s^{\nu_2-2} & & & \cdots & 0 & s^{\nu_2-\nu_1} & & \\ & & \ddots & & & & \ddots & & & & & \ddots & 0 \\ & & & s^{\nu_p-1} & 0 & & & s^{\nu_p-2} & \cdots & 0 & & 0 & s^{\nu_p-\nu_1} \end{bmatrix}.$$

In this representation, the monom s^β and the corresponding column is omitted as soon as the exponent $\beta < 0$. $\mathfrak{X}(s)$ and $X(s)$ are related by a simple permutation of the columns, i.e. there is a permutation matrix U such that $\mathfrak{X}(s) = X(s)U$. This permutation transforms the controllability matrix $\mathcal{C}(A,B)$ into $U^{-1}\mathcal{C}(A,B)$.

Although it is much simpler to explain the algorithm by performing the U transformation as above, in practice the computer would automatically perform the realization with respect to the new basis matrix $\mathfrak{X}$ and arrive at $U^{-1}AU$ and $U^{-1}B$ instead of A and B. The realization with respect to the new basis matrix is just as simple to compute as the original, yet it is in a more practical form and, by arriving at it directly, will not waste time by transforming basis matrices.

As mentioned earlier, the basis matrix $\mathfrak{X}(s)$ (as well as the basis matrix $X(s)$) has the property that every polynomial p-vector $\varphi(s) \in \mathbb{F}^p[s]$ whose i-th component has degree at most $\nu_i - 1$ can uniquely be described through $\varphi(s) = \mathfrak{X}(s)\alpha, \alpha \in \mathbb{F}^n$. It is therefore possible to identify $\varphi(s)$ with the n-vector α. We will say that α is the *coordinate vector* of $\varphi(s)$ with respect to the basis matrix $\mathfrak{X}(s)$.

Theorem 5.1 *Assume $P(s)$ has Kronecker indices $\nu_1 \geq \cdots \geq \nu_p$ and minimal indices $\mu_1, \ldots, \mu_p$ none of which equal zero. Let $L(s) = [\mathbf{l}_1, \ldots, \mathbf{l}_p]$ be a GCLD whose (i,j)-entry has degree at most $\nu_i - \mu_j$ or is zero. Assume that the matrix L_∞ is lower triangular (by Lemma 4.1) and let $d = \sum_{i=1}^p \mu_i$. Then the $n \times d$ scalar matrix whose columns form the coordinate vectors of*

$$\left[s^{\mu_1-1}\mathbf{l}_1 \ldots s\mathbf{l}_1\ \mathbf{l}_1 \mid \ldots \mid s^{\mu_p-1}\mathbf{l}_p \ldots s\mathbf{l}_p\ \mathbf{l}_p\right] \tag{5.1}$$

is after a possible permutation of the columns in column echelon form.

Proof: Immediate consequence from the fact that L_∞ is lower triangular, has nonzero diagonal elements and the specific choice of the basis matrix $\mathfrak{X}(s)$. □

As a consequence of this theorem we can immediately read out the minimal indices $\mu_1, \ldots, \mu_p$ from the pivot indices of the column echelon form of $\mathcal{C}(A, B)$. A priori it is not true that $\mathfrak{X}(s)\mathcal{C}(A, B)$ has the particular form (5.1) even if $\mathcal{C}(A, B)$ is in column echelon form. One observes however that elementary column operations on $\mathcal{C}(A, B)$ correspond to unimodular operations on $\mathfrak{X}(s)\mathcal{C}(A, B)$. By Theorem 4.3 we also know that the columns of $\mathfrak{X}(s)\mathcal{C}(A, B)$ are in the column module of $L(s)$. By the above remarks it is possible to identify p columns $[c_1, \ldots, c_p]$ from the column echelon form of $\mathcal{C}(A, B)$ such that $\mathfrak{X}(s)[c_1, \ldots, c_p]$ forms a GCLD of $P(s)$. In the sequel we make this selection process more precise.

Assume that the controllability matrix $\mathcal{C}(A, B)$ is in column echelon form. We can think of the controllability matrix as being divided into row blocks. The top row block consists of p rows and corresponds (under multiplication by $\mathfrak{X}(s)$) to coefficients of degree $\nu_i - 1$ for each respective row i. The next lower block corresponds to coefficients of degree $\nu_i - 2$. Each lower block is similarly defined. If $\nu_j - \beta < 0$ then no row corresponding to row j occurs in row block β (or any subsequent blocks). Based on this we define:

Definition 5.2 1. A column in the controllability matrix $\mathcal{C}(A, B)$ is said to "take its order in row i" if the leading coefficient occurs in a row which corresponds (under multiplication by $\mathfrak{X}(s)$) to an entry in row i of the resulting polynomial p-vector.

2. For each row i, $1 \leq i \leq p$, consider all the column vectors of the controllability matrix taking their order in row i. From this set, the column vector whose leading coefficient is lowest (in the matrix, not necessarily in value), is called the "row leader for row i".

Theorem 5.3 *If the column echelon form of $\mathcal{C}(A, B)$ has p row leaders $[c_1, \ldots, c_p]$ then $\mathfrak{X}(s)[c_1, \ldots, c_p]$ forms a GCLD of $P(s)$.*

Proof: It follows from our definition of row leaders $[c_1, \ldots, c_p]$ that

$$\deg \det \mathfrak{X}(s)[c_1, \ldots, c_p] = \sum_{i=1}^{p} \nu_i - \sum_{i=1}^{p} \mu_i = n - d.$$

Since $\mathfrak{X}(s)[c_1, \ldots, c_p]$ is a subset of the column module of $L(s)$ it follows that the columns of $\mathfrak{X}(s)[c_1, \ldots, c_p]$ generate this column module and this completes the proof. □

Remark 5.4 It can be shown and it is illustrated in an example in Section 8 that in the case of $m = p = 1$, i.e. in the situation where $P(s) = (p_1(s), p_2(s))$ the column reduction of the controllability matrix $\mathcal{C}(A, B)$ is exactly the Euclidean algorithm. The presented algorithm generalizes in this way Euclid's algorithm.

Remark 5.5 The column reduction of $\mathcal{C}(A, B)$ can be done very efficiently by iteratively computing the vectors $A^i b_j$, where b_j is the j-th column of B. (See [16, 1] for more details). Due to the very sparse structure of (A, B) the column reduction is even easier.

6 The Situation of Constant Rows

As remarked earlier, the matrix $\tilde{P}(s)$ that is used in the proof of our algorithm could have constant rows, and that poses problems when we try to realize this matrix. In this section we will deal with this case. In particular, assume that $\tilde{P}(s)$ has $0 \leq k \leq p$ constant rows ($\mu_i = 0$ for $1 \leq i \leq k$).

Similar to before, we know that $\tilde{P}(s)$ has (after possible right scalar multiplication) the form:

$$\tilde{P}(s) \;=\; \left[\begin{array}{cc|c} I_k & 0 & 0 \\ 0 & \tilde{E}(s) & \tilde{F}(s) \end{array}\right] \tag{6.1}$$

Letting $\tilde{P}_r(s) = [\tilde{E}(s) \mid \tilde{F}(s)]$, we can obtain the realization matrices $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ relative to the basis matrix $\tilde{X}(s)$ for $\tilde{P}_r(s)$. The following result can easily be shown using arguments mirroring those in the proof of Theorem 4.3:

Theorem 6.1

$$L(s) \left[\begin{array}{c} 0 \\ \tilde{X}(s) \end{array}\right] \mathcal{C}(\tilde{A}, \tilde{B}) = X(s)\mathcal{C}(A, B)$$

In analogy to Corollary 4.4 we have:

Corollary 6.2 *Let $\mu_{k+1}, \ldots, \mu_p$ be the nonzero minimal indices of $P(s)$. Then there exists an invertible matrix $W \in Gl_{mn}$ such that*

$$X(s)\mathcal{C}(A, B)W = \left[s^{\mu_{k+1}-1}\mathbf{l}_{k+1} \ldots s\mathbf{l}_{k+1}\ \mathbf{l}_{k+1} \mid \ldots \mid s^{\mu_p-1}\mathbf{l}_p \ldots s\mathbf{l}_p\ \mathbf{l}_p \;\mid\; O_{p\times(mn-d)}\right]. \tag{6.2}$$

In this representation the $[\mathbf{l}_{k+1}, \ldots, \mathbf{l}_p]$ represent $p-k$ generators of the column module of $P(s)$.

By Lemma 4.2 we know that the $\operatorname{rank}\mathcal{C}(A, B) = \sum_{i=1}^{p} \mu_i = d$. Combining this with Corollary 6.2 and Theorem 5.3 results in:

Theorem 6.3 *If $P(s)$ has k zero minimal indices then the column echelon form of $\mathcal{C}(A, B)$ has exactly $p-k$ row leaders $[c_{k+1}, \ldots, c_p]$ and the columns of $\mathfrak{X}(s)[c_{k+1}, \ldots, c_p]$ form an independent set of generators for a GCLD of $P(s)$.*

By this last theorem we will be able to compute the number, k, of nonzero minimal indices of $P(s)$ and we always will be able to identify $p-k$ 'row leaders' from the echelon form of $\mathcal{C}(A, B)$. This is very important. Otherwise, we could perform the algorithm, get p columns and think we are done, when in reality we would have selected columns that are unimodularly equivalent and ended up with a singular matrix!

The question now turns to: How do we select the remaining k columns to fill up our matrix and arrive at a GCLD?

The answer is actually quite simple. For this consider the high order coefficient matrix H of the $p \times (p-k)$ matrix $\mathfrak{X}(s)[c_{k+1}, \ldots, c_p]$. This high order coefficient matrix is a submatrix of the matrix L_∞, introduced in Lemma 4.1. The high order coefficient matrix of $P(s)$ is

assumed to be $P_h = [I_p\ 0]$. It is therefore possible to augment H with columns from P_h such that the overall matrix L_∞ becomes invertible. Correspondingly we have a way of selecting k columns from the first p columns of $P(s)$ such that $\mathfrak{x}(s)[c_{k+1}, \ldots, c_p]$ augmented by these columns results in a GCLD of $P(s)$. Simply put: For every row, i, which does not have a row leader, simply select column i from the matrix $P(s)$ to be in $L(s)$.

7 The Algorithm

We now present the algorithm of computing a GCLD in a concise form:

Step 1 We are given a full rank polynomial matrix $P(s)$.

Step 2 Check if the high order coefficient matrix P_h has the form $[I \mid 0]$. If not, then use right and left unimodular operations to bring it into this form. Keep track of any left unimodular operations in the matrix $V(s)$.

Step 3 Check if $P(s)$ has any constant rows. If $P(s)$ has κ constant rows and is in the form (3.3) then the submatrix $\begin{bmatrix} \hat{E}_2(s) \\ I_\kappa \end{bmatrix}$ of (3.3) defines κ generators of a GCLD $L(s)$. Continue the algorithm with the reduced matrix $[\, \hat{E}_1(s) \mid \hat{F}(s)\,]$ in order to find the remaining $p - \kappa$ columns of the GCLD.

Step 4 Obtain the realization matrices A and B relative to the basis matrix $\mathfrak{x}(s)$ 'by inspection'.

Step 5 Calculate the controllability matrix $\mathcal{C}(A, B)$ and column reduce it. (This may be done simultaneously to greatly improve efficiency [16, 1].)

Step 6 Pick out the 'row leaders' from the column reduced controllability matrix $\mathcal{C}(A, B)$. Multiply the 'row leaders' by $\mathfrak{x}(s)$ and place them in the GCLD.

Step 7 If there are p row leaders, then go to step 8. If there are less than p row leaders, then follow the algorithm of Section 6.

Step 8 Multiply the GCLD on the left by V^{-1} and stop.

Remark 7.1 The steps which take the most time are steps 2 and 5. Step 2 is not necessary when $P(s)$ is in the desired form. Of course, in general we will not know or can not guarantee what form a matrix will have. However, in certain applications, such as searching for observable convolutional encoders [1, 7], we may prescribe what form the matrices will have.

After having computed the GCLD, $L(s)$, there might arise the need to compute the 'controllable part' $\tilde{P}(s)$ as well. Let $\mathbf{p}_i(s)$ and $\tilde{\mathbf{p}}_i(s)$ denote the ith column of $P(s)$ and $\tilde{P}(s)$ respectively, $i = 1, \ldots, m + p$. Consider for each index i the equation

$$L(s)\tilde{\mathbf{p}}_i(s) = \mathbf{p}_i(s). \tag{7.1}$$

We can view (7.1) as a system of $n+p$ linear equations in $d+p$ unknowns. We therefore have to solve simultaneously $m+p$ systems of equations in $d+p$ unknowns. Due to the fact that the matrix L_∞ is already in lower triangular form it follows that the coefficient matrix appearing in (7.1) is already in triangular form as well. A solution of (7.1) can therefore be computed very efficiently and the method will be illustrated in the next section.

8 Examples

We have included some examples to aid in the understanding of the algorithm.

Example 8.1 First, take the case when $P(s)$ is a 1×2 matrix. In this case we are just determining the gcd of two polynomials. Notice that $P(s)$ will trivially satisfy all of the conditions unless the two polynomials have the same degree. In that case divide one into the other, and take the remainder in place of the original polynomial.

Let us work through the following example:

$$P(s) = [\, s^6+5s^5-464s^4+1123s^3-887s^2+234s+72 \qquad s^5-2s^4-342s^3+1177s^2-1170s+504 \,]$$

We get the following realization:

$$A = \begin{bmatrix} -5 & 1 & 0 & 0 & 0 & 0 \\ 464 & 0 & 1 & 0 & 0 & 0 \\ -1123 & 0 & 0 & 1 & 0 & 0 \\ 887 & 0 & 0 & 0 & 1 & 0 \\ -234 & 0 & 0 & 0 & 0 & 1 \\ -72 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 \\ -2 \\ -342 \\ 1177 \\ -1170 \\ 504 \end{bmatrix}$$

relative to the basis matrix $\mathfrak{X}(s) = [s^5\ s^4\ s^3\ s^2\ s\ 1]$. The corresponding column reduced controllability matrix is:

$$\mathcal{C}(A,B) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 & 0 \\ -342 & \frac{-65}{3} & 1 & 0 & 0 & 0 \\ 1177 & \frac{221}{3} & -19 & 0 & 0 & 0 \\ -1170 & \frac{-220}{3} & 23 & 0 & 0 & 0 \\ 504 & 32 & -12 & 0 & 0 & 0 \end{bmatrix}$$

Since there is only one row of $\mathfrak{X}(s)\mathcal{C}(A,B)$, the row leader must be the rightmost nonzero column. Hence the GCLD is $s^3 - 19s^2 + 23s - 12$.

Notice that the first column of the above matrix corresponds with polynomial of lesser degree from our original matrix. The second column corresponds with the 'first remainder' that one obtains when applying the Euclidean algorithm to the two polynomials in our matrix. The third column corresponds with the 'second remainder', and also the last nonzero one, of the Euclidean algorithm. Because of this, our algorithm can be seen as an extension of the Euclidean algorithm to matrices.

Example 8.2 Now let us look at a more nontrivial example.

$$P(s) := \begin{bmatrix} s^5 & s^4+s^2 & s^4+2\,s^2 \\ s & s^3+s^2+s+1 & 2\,s+3 \end{bmatrix}$$

Here $\nu_1 = 5, \nu_2 = 3$ and the realization matrices are

$$A = \begin{bmatrix} 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 2 \\ 3 \\ 0 \\ 0 \end{bmatrix}$$

relative to the basis matrix

$$\mathfrak{X}(s) = \begin{bmatrix} s^4 & 0 & s^3 & 0 & s^2 & 0 & s & 1 \\ 0 & s^2 & 0 & s & 0 & 1 & 0 & 0 \end{bmatrix}.$$

The column reduced controllability matrix is

$$\mathcal{C}(A,B) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We see that columns 2 and 3 take their order in row 2, while column 1 is the only column taking its order in row 1. Hence column 1 is the 'row leader' for row 1 and column 3 is the 'row leader' for row 2. It follows that $\mu_1 = 1$ and $\mu_2 = 2$. As an independent verification, we can also see directly from $\mathfrak{X}(s)\mathcal{C}(A,B)$ that column 2 is just $s-1$ times column 3 and hence they are dependent.

$$\mathfrak{X}(s)\mathcal{C}(A,B) = \begin{bmatrix} s^4 & s^3-s^2 & s^2 & 0 & 0 & 0 & 0 & 0 \\ 1 & s^2-1 & s+1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

so the GCLD is

$$L(s) = \begin{bmatrix} s^4 & s^2 \\ 1 & s+1 \end{bmatrix}.$$

We can now also easily compute $\tilde{P}(s)$ by solving the following linear system of equations:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \\ e_1 & e_2 & e_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

This corresponds to the equation $L(s)\tilde{P}(s) = P(s)$, where $\tilde{P}(s)$ is represented by the matrix:

$$\begin{bmatrix} a_1 s + b_1 & a_2 s + b_2 & a_3 s + b_3 \\ c_1 s^2 + d_1 s + e_1 & c_2 s^2 + d_2 s + e_2 & c_3 s^2 + d_3 s + e_3 \end{bmatrix}.$$

The left-hand matrix in the above equation comes easily from the column reduced controllability matrix. It consists of the 'row leaders' plus 'shifts' of the row leaders. To be precise, for each i, the row leader for column i occurs, along with μ_i upward 'shifts' of the row leader. Note that this necessitates adding another row block to the top of the scalar matrix.

The right-hand matrix is simply the coefficients of the matrix $P(s)$ with respect to the 'augmented basis matrix' $\left[\begin{array}{cc|c} s^5 & 0 & \\ 0 & s^3 & \mathfrak{X}(s) \end{array}\right]$.

Not only is the left-hand matrix easily constructed, but it will be lower diagonal (up to column permutations) so that the above system can be solved instantaneously! The resulting matrix $\tilde{P}(s)$ can now be stated:

$$\tilde{P}(s) = \begin{bmatrix} s & 0 & 1 \\ 0 & s^2 + 1 & 2 \end{bmatrix}.$$

References

[1] B.M. Allen and J. Rosenthal. Analysis of convolutional encoders via generalized sylvester matrices and state space realization. In *Proc. of the 34-th Allerton Conference on Communication, Control, and Computing*, pages 893–902, 1996.

[2] B.M. Allen and J. Rosenthal. Analyzing convolutional encoders using realization theory. In *Proceedings of the 1997 IEEE International Symposium on Information Theory*, page 287, Ulm, Germany, 1997.

[3] A.C. Antoulas. On recursiveness and related topics in linear systems. *IEEE Trans. Automat. Contr.*, AC-31(12):1121–1135, 1986.

[4] J. A. Ball, I. Gohberg, and L. Rodman. *Interpolation of Rational Matrix Functions.* Birkhäuser Verlag, Basel-Berlin-Boston, 1990.

[5] S. Barnett. *Polynomials and linear control systems.* M. Dekker, New York, 1983.

[6] R. R. Bitmead, S. Y. Kung, B. D. O. Anderson, and T. Kailath. Greatest common divisors via generalized Sylvester and Bézout matrices. *IEEE Trans. Automat. Control*, AC-23:1043–1047, 1978.

[7] A. Dholakia. *Introduction to Convolutional Codes with Applications.* Kluwer Academic Publishers, 1994.

[8] G. D. Forney. Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM J. Control Optim.*, 13(3):493–520, 1975.

[9] Paul A. Fuhrmann. A matrix euclidean algorithm and matrix continued fraction expansions. *Systems and Control Letters*, 3:263–271, 1983.

[10] U. Helmke, J. Rosenthal, and J. M. Schumacher. A controllability test for general first-order representations. *Automatica*, 33(2):193–201, 1997.

[11] T. Kailath. *Linear Systems.* Prentice-Hall, Englewood Cliffs, N.J., 1980.

[12] S.-Y. Kung, T. Kailath, and M. Morf. A generalized resultant matrix for polynomial matrices. In *Proc. IEEE Conf. on Decision and Control (Florida)*, pages 892–895, 1976.

[13] C. C. MacDuffee. Some applications of matrices in the theory of equations. *American Mathematical Monthly*, 57:154–161, 1950.

[14] M. S. Ravi, J. Rosenthal, and J. M. Schumacher. Homogeneous behaviors. *Math. Contr., Sign., and Syst.*, 10:61–75, 1997.

[15] J. Rosenthal and J. M. Schumacher. Realization by inspection. *IEEE Trans. Automat. Contr.*, AC-42(9):1257–1263, 1997.

[16] E. D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems.* Springer Verlag, New York, 1990.

[17] J. C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Trans. Automat. Control*, AC-36(3):259–294, 1991.

[18] W. A. Wolovich. *Linear Multivariable Systems*, volume 11 of *Appl. Math. Sc.* Springer Verlag, New York, 1974.