Embedding polynomial matrices of one variable

van der Put, Marius; Tsang, Fai Lung

Published in:
Indagationes mathematicae-New series

Publication date:
2008

Indag. Mathem., N.S., 19 (4), 643–648 December, 2008

Embedding polynomial matrices of one variable

by Marius van der Put[a] and Fai Lung Tsang[b]

[a] *Institute of Mathematics and Computer Science, University of Groningen, P.O. Box 407, 9700 AK Groningen, The Netherlands*
[b] *Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

Communicated by Prof. M.S. Keane

ABSTRACT

A non square matrix with coefficients in $K[z]$ can (if a condition on its minors is satisfied) be embedded into a square matrix with determinant 1. Finding theoretically and in an algorithmic way an embedding of small degree is solved by a construction with vector bundles on the projective line over K.

1. INTRODUCTION

Let K be any field and z a variable. Let a, n denote integers such that $1 \leqslant a < n$ and let P be an $(a \times n)$-matrix with coefficients in $K[z]$. It is an elementary and well-known fact that there exists an $((n-a) \times n)$-matrix Q with coefficients in $K[z]$, such that the $(n \times n)$-matrix $\binom{P}{Q}$ is invertible (i.e., its determinant lies in K^*) if and only if the $(a \times a)$-minors of P have no common zero. The *problem*, recently posed to us by B. Ćurgus and A. Dijksma, concerns the minimal degree of a matrix Q with this property. A rough answer is: *If the degree of P is $D \geqslant 1$, then Q exists with degree $\leqslant D - 1$.*

Using some properties of vector bundles on the projective line $\mathbb{P}^1$ over K we will prove this and develop an algorithm for Q consisting of only a few steps of linear algebra over K. The result turned out to be known and the first reference for it is (probably) [2]. The proof there is a computation with matrix polynomials (pencils).

MSC: 15A33, 14F05, 46M20
Key words and phrases: Polynomial matrix, Vector bundles on the projective line, Forney sequence, Minimal basis, Saturated module

E-mails: M.van.der.Put@math.rug.nl (M. van der Put), eefl@ust.hk (F.L. Tsang).

In [1] this proof is written out and expanded into an algorithm which is numerically stable.

A first translation of the problem is as follows. Let V be a vector space over K of dimension n. Then $V[z] := V \bigotimes_K K[z]$ is a free $K[z]$-module of rank n. The elements of $V[z]$ can be written as finite sums $\sum_i v_i z^i$ with coefficients in V. The degree of a non zero element is $\max\{i \mid v_i \neq 0\}$.

After choosing a basis $b_1, \ldots, b_n$ of V, one can identify V with K^n and $v(\in V[z])$ with a n-tuple of polynomials $(p_1(z), \ldots, p_n(z))$. The degree of v is equal to $\max(\deg p_i(z))$. This observation shows that no choice of a basis of V is needed for the sequel.

The image of P, seen as a $K[z]$-linear map $W[z] \to V[z]$, where W is a vector space over K of dimension a, is a submodule M of $V[z]$ (again the choices of bases for V and W are immaterial). The condition on P is equivalent to the condition '$M \subset V[z]$ is *saturated*', i.e., if $f \cdot r \in M$, $f \in K[z]$, $f \neq 0$, $r \in V[z]$, then $r \in M$. The required Q corresponds to a *complement* N of M, i.e., a submodule N of $V[z]$ such that $M \oplus N = V[z]$. The answer Proposition 1.1 to the above problem is more precise since d, i.e., the maximum of the Forney sequence of M (see Section 2), or, equivalently, the minimal number such that M is generated by elements of degree $\leqslant d$, satisfies $d \leqslant D$ where D is the degree of the matrix P.

Proposition 1.1. *A saturated submodule M generated by elements of degree $\leqslant d$ has a complement N generated by elements of degree* $\leqslant \max\{d-1, 0\}$.

2. FORNEY SEQUENCES AND VECTOR BUNDLES ON $\mathbb{P}^1$

Let $M \subset V[z]$ be a submodule of rank $a > 0$. We recall that M is a free $K[z]$-module. For any integer k we put $M_k = \{m \in M \mid \deg m \leqslant k\}$ (and thus $M_k = 0$ if $k < 0$). For $k \geqslant 0$, the linear map $top_k : M_k \to V$ assigns to every element in M_k the coefficient of z^k. The image of the injection $M_k/M_{k-1} \to V$ is written as W_k. One has inclusions $W_{-1} := 0 \subset W_0 \subset W_1 \subset \cdots$. The *Forney sequence* $k_1 \leqslant k_2 \leqslant \cdots \leqslant k_a$ of M is the sequence of the integers k such that $W_k \neq W_{k-1}$, repeated with multiplicity $\dim W_k/W_{k-1}$. A basis $m_1, \ldots, m_a$ for M (i.e., free generators of the $K[z]$-modules M) is called *minimal* if $\sum_{i=1}^{a} \deg m_i$ is minimal among the set of all bases.

A minimal basis can be obtained as follows. Let $a_1 < a_2 < \cdots < a_r$ denote the integers k such that $W_k \neq W_{k-1}$. For each a_i one takes elements in M_{a_i} that are mapped by $M_{a_i} \overset{top_{a_i}}{\to} W_{a_i} \to W_{a_i}/W_{a_i-1}$ to a basis of W_{a_i}/W_{a_i-1}. The set of all these elements is a minimal basis. The statements in the following lemma are easily verified.

Lemma 2.1.

1. *Every minimal basis is obtained as above. Moreover, the degrees of the elements of a minimal basis form the Forney sequence of M.*
2. *Let $m_1, \ldots, m_a$ be a minimal basis. The degree of an element $\sum f_i m_i$ (with all $f_i \in K[z]$) is equal to* $\max\{\deg f_i + \deg m_i\}$.

3. *M_k consists of the elements $\sum f_i m_i$ with $f_i \in K[z]_{k-\deg m_i}$ (i.e., the elements of $K[z]$ of degree $\leqslant k - \deg m_i$). This K-vector space has dimension $\sum_{i,\deg m_i \leqslant k}(k - \deg m_i + 1)$.*

We recall some properties of vector bundles on $\mathbb{P}^1$ and refer for more details to Hartshorne [3] (especially Chapter V, Section 2, Exercise 2.6). The structure sheaf on $\mathbb{P}^1$ is denoted by $\mathcal{O}$. For any divisor $D = \sum n_i[p_i]$ on $\mathbb{P}^1$ we write $\mathcal{O}(D)$ for the corresponding line bundle. The degree of $\mathcal{O}(D)$ is the degree of D and thus equal to $\sum n_i \cdot \deg p_i$. We identify the standard line bundle $\mathcal{O}(1)$ of degree one with $\mathcal{O}([\infty])$. A vector bundle of rank a on $\mathbb{P}^1$ is a coherent sheaf, locally isomorphic to $\mathcal{O}^a$ (i.e., the direct sum of a copies of $\mathcal{O}$). The classical result, which carries the name of Grothendieck, states that every vector bundle decomposes as a direct sum of line bundles and every line bundle is isomorphic to $\mathcal{O}(k)$ for a unique integer k. For a vector bundle (or any coherent sheaf) $\mathcal{F}$ we write $H^i(\mathcal{F})$, $i = 0, 1$, for the cohomology groups on $\mathbb{P}^1$. Now $\dim H^0(\mathcal{O}(k))$ is zero if $k < 0$ and equal to $k + 1$ if $k \geqslant 0$. Further, $\dim H^1(\mathcal{O}(k))$ is zero for $k \geqslant -1$ and is $-k - 1$ for $k < -1$. For any vector bundle (or coherent sheaf) $\mathcal{F}$ and any integer k we write $\mathcal{F}(k) = \mathcal{F} \otimes \mathcal{O}(k)$. In particular, $\mathcal{O}(\ell)(k) = \mathcal{O}(k + \ell)$. For any vector bundle (or coherent sheaf) $\mathcal{F}$ one defines the vector bundle (or coherent sheaf) $\mathcal{F}^* = \mathcal{H}om_{\mathcal{O}}(\mathcal{F}, \mathcal{O})$. We note that $\mathcal{O}(k)^* = \mathcal{O}(-k)$.

We associate to $K[z]$ (provided with its degree function) the sheaf $\mathcal{O}$ and to $V[z]$ (with $\dim_K V = n$ and provided with its degree function) the sheaf $V \bigotimes_K \mathcal{O} \cong \mathcal{O}^n$. One observes that $H^0(V \bigotimes_K \mathcal{O}(d)) := H^0(\mathbb{P}^1, V \bigotimes_K \mathcal{O}(d))$ is equal to $\{v \in V[z] \mid \deg v \leqslant d\}$ and so the degree function on $V[z]$ is recovered from the sheaf $V \bigotimes_K \mathcal{O}$. Write $\mathbb{A}^1 = \mathbb{P}^1 \backslash \{\infty\}$. Then one has $H^0(\mathbb{A}^1, V \bigotimes_K \mathcal{O}) = \bigcup_{d \geqslant 0} H^0(V \bigotimes_K \mathcal{O}(d)) = V[z]$.

Let $m \in V[z]$ be an element of degree d. Then $m \in H^0(V \otimes \mathcal{O}(d))$ and there is a unique morphism $f : \mathcal{O} \to V \otimes \mathcal{O}(d)$ of vector bundles such that f maps the global section 1 to the global section m. By taking the tensor product with $\mathcal{O}(-d)$, one obtains a morphism $f(-d) : \mathcal{O}(-d) \to V \otimes \mathcal{O}$. The image is isomorphic to $\mathcal{O}(-d)$ and will be denoted by $\mathcal{O}(-d)m \subset V \otimes \mathcal{O}$.

Lemma 2.2. *Let M be a submodule of $V[z]$ of rank a. There is a unique subbundle $\mathcal{M}$ of $V \bigotimes_K \mathcal{O} \cong \mathcal{O}^n$ such that $H^0(\mathbb{A}^1, \mathcal{M}) = M$ and for all integers $k \geqslant 0$, $H^0(\mathcal{M}(k)) = \{m \in M \mid \deg m \leqslant k\}$.*

Further, $\mathcal{M} \cong \mathcal{O}(-k_1) \oplus \cdots \oplus \mathcal{O}(-k_a)$, where $k_1 \leqslant \cdots \leqslant k_a$ is the Forney sequence of M. Moreover, M is saturated if and only if the quotient sheaf $\mathcal{O}^n/\mathcal{M}$ is a vector bundle.

Proof. That $\mathcal{M}$ is unique easily follows from the condition: "for all integers $k \geqslant 0$, $H^0(\mathcal{M}(k)) = \{m \in M \mid \deg m \leqslant k\}$". For its existence, we consider a minimal basis $m_1, \ldots, m_a$ for M such that $\deg m_i = k_i$ for $i = 1, \ldots, a$. Using Lemma 2.1, part (3), one sees that the required properties for $\mathcal{M}$ are fulfilled precisely for the subsheaf $\mathcal{M} = \bigoplus_{i=1}^a \mathcal{O}(-k_i)m_i$ of $\mathcal{O}^n$. The relation with the Forney sequence is evident.

Write $\mathcal{N} := (V \otimes \mathcal{O})/\mathcal{M}$. The local ring $\mathcal{O}_\infty$ of $\mathcal{O}$ at ∞ is $K[t]_{(t)}$ and its completion is $K[[t]]$, where $t = z^{-1}$. For the completion of the stalk of $\mathcal{N}_\infty$ of $\mathcal{N}$ at ∞, one has the exact sequence

$$0 \to \bigoplus_{i=1}^{a} t^{k_i} K[[t]]m_i \to V \otimes K[[t]] \to \mathcal{N}_\infty \otimes_{K[t]_{(t)}} K[[t]] \to 0.$$

Now, $t^{k_i} m_i = t^{k_i}(v_i z^{k_i} + *z^{k_i-1} + \cdots) = v_i + *t + \cdots$, where $v_i = top_{k_i} m_i$. By construction, the elements $v_1, \ldots, v_a$ are linearly independent. It follows that $\mathcal{N}_\infty \otimes_{K[t]_{(t)}} K[[t]]$ is isomorphic to $K[[t]]^{n-a}$ and thus $\mathcal{N}_\infty \cong \mathcal{O}_\infty^{n-a}$. Hence $\mathcal{N}_\infty$ is free over $\mathcal{O}_\infty$. For a point $p \neq \infty$ the stalk $\mathcal{N}_p$ is equal to $V[z]_{(p)}/M_{(p)}$, where (p) denotes the maximal ideal of $K[z]$ associated to the point p. We conclude that $\mathcal{N}$ is locally free (i.e., is a vector bundle) if and only if M is saturated. Indeed, M is saturated if and only if $V[z]_{(p)}/M_{(p)}$ is free for all maximal ideals (p) of $K[z]$.

3. THE PROOF OF PROPOSITION 1.1

The saturated submodule M corresponds to a subsheaf $\mathcal{M}$ of $V \otimes \mathcal{O} \cong \mathcal{O}^n$ on $\mathbb{P}^1$ such that $\mathcal{N}$ defined by the exactness of the sequence

$$0 \to \mathcal{M} \to V \otimes \mathcal{O} \to \mathcal{N} \to 0 \text{ is a vector bundle.}$$

Let $0 \leqslant k_1 \leqslant \cdots \leqslant k_a$ be the Forney sequence of M, then $\mathcal{M} \cong \mathcal{O}(-k_1) \oplus \cdots \oplus \mathcal{O}(-k_a)$. We suppose that $k_a \geqslant 1$ (otherwise the result that we want to prove is trivial). Further, as we know, $\mathcal{N}$ is a direct sum of line bundles $\mathcal{O}(\ell_i), i = 1, \ldots, b$, with $\ell_1 \leqslant \cdots \leqslant \ell_b$ and $a + b = n$. We write $\mathcal{N} = \mathcal{O}(\ell_1)e_1 \oplus \cdots \oplus \mathcal{O}(\ell_b)e_b$, where the $e_1, \ldots, e_b$ are written to keep track of the factors. Using the earlier notation, e_i is a global section of $\mathcal{N}(-\ell_i)$. We note that the sections of $\mathcal{N}$ above $\mathbb{A}^1$ is just $K[z]e_1 \oplus \cdots \oplus K[z]e_b$.

Tensoring the above exact sequence with $\mathcal{O}(-1)$, taking cohomology and using that $H^1(\mathcal{O}(-1)) = 0$ one finds that $H^1(\mathcal{N}(-1)) = 0$. This implies that $\ell_1 \geqslant 0$. The exactness of

$$0 \to \mathcal{M}(k_a - 1) \to V \otimes \mathcal{O}(k_a - 1) \to \mathcal{N}(k_a - 1) \to 0$$

and the vanishing of $H^1(\mathcal{M}(k_a - 1))$ induces a surjective map $V \otimes K[z]_{k_a-1} = H^0(V \otimes \mathcal{O}(k_a - 1)) \to H^0(\mathcal{N}(k_a - 1))$. The last vector space contains the elements $e_1, \ldots, e_b \in H^0(\mathcal{N})$. Take elements $f_1, \ldots, f_b \in V \otimes K[z]_{k_a-1}$ which are mapped to $e_1, \ldots, e_b$, and let N be the submodule of $V[z]$ generated by $f_1, \ldots, f_b$. Taking the sections above $\mathbb{A}^1$ of the last exact sequence of vector bundles yields the exact sequence of $K[z]$-modules

$$0 \to M \to V[z] \to \bigoplus_{i=1}^{b} K[z]e_i \to 0.$$

This clearly implies that $M \oplus N = V[z]$.

We use the notation of the proof of Proposition 1.1. The $K[z]$-module $Q := H^0(\mathbb{A}^1, \mathcal{N})$ has the free basis $e_1, \ldots, e_b \in H^0(\mathcal{N})$ and there is an exact sequence $0 \to M \to V[z] \to Q \to 0$. If the elements $e_1, \ldots, e_b \in H^0(\mathcal{N}) \subset Q$ are known, then trivial linear algebra produces elements $f_i \in V \otimes K[z]_{k_a-1}$ with image e_i for $i = 1, \ldots, b$. Further $f_1, \ldots, f_b$ is a free basis of a complement N of M, having the required property of the degrees. Therefore we have to produce an algorithm computing $e_1, \ldots, e_b$.

For any $K[z]$-module A we write $A^* = \mathrm{Hom}_{K[z]}(A, K[z])$. The module Q^* has a dual (free) basis $L_1, \ldots, L_b$ given by $L_i(e_j) = \delta_{ij}$. From this dual basis one can recover $e_1, \ldots, e_b$ by linear algebra over $K(z)$.

Recall that $\mathcal{N} = \bigoplus_{i=1}^b \mathcal{O}(\ell_i)e_i$ and consider, for $i = 1, \ldots, b$, the morphism

$$L_i : \mathcal{N}(-\ell_i) = \bigoplus_j \mathcal{O}(\ell_j - \ell_i)e_j \xrightarrow{\text{projection}} \mathcal{O}e_i \overset{e_i \mapsto 1}{\to} \mathcal{O}.$$

Then $L_i \in H^0(\mathcal{N}^*(\ell_i)) \subset H^0(\mathbb{A}^1, \mathcal{N}^*) = Q^*$. Since $L_i(e_j) = \delta_{ij}$ holds, $\{L_i\}$ is the dual basis. Further $\mathcal{N}^* = \bigoplus_i \mathcal{O}(-\ell_i)L_i$.

Dualizing the sequence $0 \to M \overset{i}{\to} V[z] \to Q \to 0$, with respect to $\mathrm{Hom}_{K[z]}(-, K[z])$, yields the exact sequence $0 \to Q^* \to V^*[z] \overset{i^*}{\to} M^* \to 0$. One observes that $Q^* = \ker i^*$, that $L_1, \ldots, L_b$ is a minimal basis for $Q^* \subset V^*[z]$ and that $\ell_1, \ldots, \ell_b$ is the Forney sequence of Q^*.

On the other hand, if $L_1, \ldots, L_b$ for Q^* is a minimal basis, then the elements $e_i \in Q$, defined by $L_i(e_j) = \delta_{i,j}$ (i.e., the dual basis) satisfy $\mathcal{N} = \bigoplus_{i=1}^b \mathcal{O}(\ell_i)e_i$. Thus we have to compute a minimal basis $L_1, \ldots, L_b$ for Q^*.

In the above we have identified the dual $V[z]^* := \mathrm{Hom}_{K[z]}(V[z], K[z])$ with $V^*[z]$ in the following way: $(\sum w_n^* z^n)(\sum v_m z^m) = \sum_{n,m \geqslant 0} w_n^*(v_m) z^{n+m}$.

Let $m_1, \ldots, m_a$ denote generators for M. Then $Q^* \subset V^*[z]$ consists of the elements $w \in V^*[z]$ with $w(m_1) = \cdots = w(m_a) = 0$. The vector spaces $Q_t^* = \{w \in Q^* \mid \deg w \leqslant t\}$ can be computed by linear algebra. One has $\sum_{i=1}^b \ell_i = \sum_{j=1}^a k_j$, since the degree of $V \otimes \mathcal{O}$ is the sum of the degrees of $\mathcal{M}$ and $\mathcal{N}$. In particular, the highest ℓ_b is $\leqslant \sum_{j=1}^a k_j$. Moreover, the computation of Q_t^* is only needed for $t \leqslant \ell_b$. As explained in the beginning of Section 2, the computation of the 'top-coefficients' $top_t(Q_t^*) \subset V^*$ leads to a minimal basis $L_1, \ldots, L_b$.

Remarks 4.1.

1. One might try to compute suitable $f_1, \ldots, f_b$ without the knowledge of the $e_1, \ldots, e_b$. This turns out to produce, even for small values of a and n, a highly complex system of polynomial equations for the many unknown coefficients of the $f_1, \ldots, f_b$.
2. For the algorithm one has to know a free basis $m_1, \ldots, m_a$ of M. This need not be a minimal basis.

3. If one identifies V with K^n and takes on K^n the canonical K-bilinear inner product $(-,-)$, then one has identified V^* with V. This yields an inner product on $V[z]$ by

$$\left(\sum v_n z^n, \sum v'_m z^m\right) = \sum_{n,m \geqslant 0} (v_n, v'_m) z^{n+m},$$

and an identification of $V[z]^*$ with $V[z]$. This is possibly a handy way to formulate the linear algebra of the computation of the spaces Q_t^*.

4. For a given saturated submodule $M \subset V[z]$ with Forney sequence (k_j), one might expect that the numbers $0 \leqslant \ell_1 \leqslant \cdots \leqslant \ell_b$ can be deduced from this Forney sequence. The following example indicates that just any sequence (ℓ_i) satisfying $\sum \ell_i = \sum k_j$ might be possible.

Example. Let M be the kernel of the map $(Ke_0 + \cdots + Ke_a)[z] \to K[z]$, given by $e_i \mapsto z^i$. Then M has minimal basis $ze_0 - e_1, ze_1 - e_2, \ldots, ze_{a-1} - e_a$ and its Forney sequence is $1, \ldots, 1$. The corresponding exact sequence $0 \to \mathcal{M} \to \mathcal{O}^{a+1} \to \mathcal{N} \to 0$ shows that $\mathcal{N} \cong \mathcal{O}(a)$.

It can be shown that a saturated submodule M with Forney sequence $0, \ldots, 0, 1, \ldots, 1$ is a direct sum of copies of the above example and trivial factors. This implies, at least in this case, that any sequence $0 \leqslant \ell_1 \leqslant \cdots \leqslant \ell_b$ with $\sum \ell_i$ equal to the number of 1's, is possible.

5. In the special case that all ℓ_i are distinct, the minimal basis $L_1, \ldots, L_b$ of Q^* is essentially unique and the same holds for $e_1, \ldots, e_b \in Q$. The only freedom in the choice of the complement N of M (with the degree restriction) is the choice of the $f_1, \ldots, f_b \in V[z]_{k_a - 1}$ with images $e_1, \ldots, e_b$.

REFERENCES

[1] Beelen T., van Dooren P. – A pencil approach for embedding a polynomial matrix into a unimodular matrix, SIAM J. Matrix Anal. Appl. **9** (1) (1988) 77–89.

[2] Eising R. – Polynomial matrices and feedback, IEEE Trans. Automat. Control **30** (10) (1985) 1022–1025.

[3] Hartshorne R. – Algebraic Geometry, Grad. Texts in Math., vol. 52, Springer-Verlag, New York–Heidelberg, 1977.

(Received July 2008)