Computing generators for linearly recurrent matrix sequences

November 8, 2017

1 Computing the canonical generator of a linearly recurrent matrix sequence

We first present the notion of linear recurrence for sequences of matrices over a field $\mathbb{K}$, which extends the well-known notion for sequences in $\mathbb{K}^{\mathbb{N}}$.

Definition 1.1 ([7, Sec. 3]). *Let $\mathcal{S} = (S_k)_{k\in\mathbb{N}} \subset \mathbb{K}^{m\times n}$ be a matrix sequence. Then,*

- *a polynomial $p = \sum_{0\le k\le d} p_k X^k \in \mathbb{K}[X]$ is said to be a* scalar relation *for $\mathcal{S}$ if $\sum_{0\le k\le d} p_k S_{\delta+k} = 0$ holds for all $\delta \ge 0$;*
- *a polynomial vector $\mathbf{p} = \sum_{0\le k\le d} p_k X^k \in \mathbb{K}[X]^{1\times m}$ is said to be a* (left, vector) relation *for $\mathcal{S}$ if $\sum_{0\le k\le d} p_k S_{\delta+k} = 0$ holds for all $\delta \ge 0$;*
- *$\mathcal{S}$ is said to be* linearly recurrent *if there exists a nontrivial scalar relation for $\mathcal{S}$.*

For designing efficient algorithms it will be useful to rely on operations on polynomials or truncated series, hence the following characterization of vector relations.

Lemma 1.2. *Consider a matrix sequence $\mathcal{S} = (S_k)_{k\in\mathbb{N}} \subset \mathbb{K}^{m\times n}$ and its generating series $\mathbf{S} = \sum_{k\ge 0} S_k/X^{k+1} \in \mathbb{K}[\![X^{-1}]\!]^{m\times n}$. Then, $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ is a vector relation for $\mathcal{S}$ if and only if the entries of $\mathbf{q} = \mathbf{p}\mathbf{S}$ are in $\mathbb{K}[X]$; furthermore, in this case, $\deg(\mathbf{q}) < \deg(\mathbf{p})$.*

Proof. Let $\mathbf{p} = \sum_{0\le k\le d} p_k X^k$. For $\delta \ge 0$, the coefficient of $\mathbf{q}$ of degree $-\delta - 1 < 0$ is $\sum_{0\le k\le d} p_k S_{k+\delta}$. Hence the equivalence, by definition of a relation. The degree comparison is clear since $\mathbf{S}$ has only terms of (strictly) negative degree. □

Concerning the algebraic structure of the set of vector relations, we have the following basic result, which can be found for example in [12, 7, 10].

Lemma 1.3. *The sequence $\mathcal{S}$ is linearly recurrent if and only if the set of left vector relations for $\mathcal{S}$ is a $\mathbb{K}[X]$-submodule of $\mathbb{K}[X]^{1\times m}$ of rank m.*

Proof. The set of vector relations for $\mathcal{S}$ is a $\mathbb{K}[X]$-submodule of $\mathbb{K}[X]^{1\times m}$, and hence is free of rank at most m [2, Chap. 12].

If $\mathcal{S}$ is linearly recurrent, let $p \in \mathbb{K}[X]$ be a nontrivial scalar relation for $\mathcal{S}$. Then each vector $[0 \ \cdots \ 0 \ p \ 0 \ \cdots \ 0]$ with p at index $1 \le i \le m$ is a vector relation for $\mathcal{S}$, hence $\mathcal{S}$ has rank m. Conversely, if $\mathcal{S}$ has rank m, then it has a basis with m vectors, which form a matrix in $\mathbb{K}[X]^{m\times m}$; the determinant of this matrix is a nontrivial scalar relation for $\mathcal{S}$. □

Note however that a matrix sequence may admit nontrivial vector relations and have no scalar relation (and therefore not be linearly recurrent with the present definition); in this case the module of vector relations has rank less than m.

Definition 1.4. *Let $\mathcal{S} \subset \mathbb{K}^{m\times n}$ be linearly recurrent. A* (left) matrix generator *for $\mathcal{S}$ is a matrix in $\mathbb{K}[X]^{m\times m}$ whose rows form a basis of the module of left vector relations for $\mathcal{S}$. This basis is said to be*

- minimal *if the matrix is row reduced [14, 6];*
- canonical *if the matrix is in Popov form [8, 6].*

Note that the canonical generator is also a minimal generator; furthermore, all matrix generators $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ for $\mathcal{S}$ have the same determinantal degree $\deg(\det(\mathbf{P}))$, which we denote by $\Delta(\mathcal{S})$. We now show that minimal matrix generators are denominators in some irreducible fraction description of the generating series of the sequence. This is a direct consequence of Lemmas 1.2 and 1.3 and of basic properties of polynomial matrices.

Corollary 1.5. *A matrix sequence $\mathcal{S} = (S_k)_{k\in\mathbb{N}} \subset \mathbb{K}^{m\times n}$ is linearly recurrent if and only if its generating series $\mathbf{S} = \sum_{k\ge 0} S_k/X^{k+1} \in \mathbb{K}[\![X^{-1}]\!]^{m\times n}$ can be written as a matrix fraction $\mathbf{S} = \mathbf{P}^{-1}\mathbf{Q}$ where $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ is nonsingular and $\mathbf{Q} \in \mathbb{K}[X]^{m\times n}$. In this case, we have $\operatorname{rdeg}(\mathbf{Q}) < \operatorname{rdeg}(\mathbf{P})$ and $\deg(\det(\mathbf{P})) \ge \Delta(\mathcal{S})$, and $\mathbf{P}$ is a matrix generator of $\mathcal{S}$ if and only if $\deg(\det(\mathbf{P})) = \Delta(\mathcal{S})$ or, equivalently, the fraction $\mathbf{P}^{-1}\mathbf{Q}$ is irreducible (that is, $\mathbf{U}\mathbf{P}+\mathbf{V}\mathbf{Q} = \mathbf{I}$ for some polynomial matrices $\mathbf{U}$ and $\mathbf{V}$).*

We remark that we may also consider vector relations operating on the right: in particular, Lemma 1.2 shows that if the sequence is linearly recurrent then these right relations form a submodule of $\mathbb{K}[X]^{n\times 1}$ of rank n. Thus, a linearly recurrent sequence also admits a right canonical generator.

Now, we focus on our algorithmic problem: given a linearly recurrent sequence, find a minimal matrix generator. We assume the availability of bounds (d_ℓ, d_r) on the degrees of the left and right canonical generators, which allow us to control the number of terms of the sequence we will access during the algorithm. Since taking the Popov form of a reduced matrix does not change the degree, any left minimal matrix generator $\mathbf{P}$ has the same degree $\deg(\mathbf{P})$ as the left canonical generator: thus, d_ℓ is also a bound on the degree of any left minimal generator. The same remark holds for d_r and right minimal generators. (These bounds d_ℓ, d_r are the same as γ_1, γ_2 in [10, Def. 4.6 and 4.7]; see also δ_l, δ_r in [13, Sec. 4.2].)

Lemma 1.6. *Let $\mathcal{S} = (S_k)_k \subset \mathbb{K}^{m\times n}$ be linearly recurrent and let $d_r \in \mathbb{N}$ be such that the right canonical matrix generator of $\mathcal{S}$ has degree at most d_r. Then, $\mathbf{p} = \sum_{0\le k\le d} p_k X^k \in \mathbb{K}[X]^{1\times m}$ is a left relation for $\mathcal{S}$ if and only if $\sum_{0\le k\le d} p_k S_{\delta+k} = 0$ holds for $\delta \in \{0, \dots, d_r - 1\}$.*

Proof. Since the right canonical generator $\mathbf{P} \in \mathbb{K}[X]^{n\times n}$ is in column Popov form, we have $\mathbf{P} = \mathbf{L}\mathrm{Diag}(X^{t_1}, \dots, X^{t_n}) - \mathbf{Q}$ where $\mathrm{cdeg}(\mathbf{Q}) < \mathrm{cdeg}(\mathbf{P}) = (t_1, \dots, t_n)$ componentwise and $\mathbf{L} \in \mathbb{K}^{n\times n}$ is unit upper triangular. We define the matrix $\mathbf{U} = \mathrm{Diag}(X^{d_r - t_1}, \dots, X^{d_r - t_n})\mathbf{L}^{-1}$, which is in $\mathbb{K}[X]^{n\times n}$ since $d_r \geq \deg(\mathbf{P}) = \max_j t_j$. Then, the columns of the right multiple $\mathbf{PU} = X^{d_r}\mathbf{I}_n - \mathbf{QU}$ are right relations for $\mathcal{S}$, and we have $\deg(\mathbf{QU}) < d_r$. As a consequence, writing $\mathbf{QU} = \sum_{0\leq k<d_r} Q_k X^k$, we have $S_{d_r+\delta} = \sum_{0\leq k<d_r} S_{k+\delta} Q_k$ for all $\delta \geq 0$.

Assuming that $\sum_{0\leq k\leq d} p_k S_{\delta+k} = 0$ holds for all $\delta \in \{0, \dots, d_r - 1\}$, we prove by induction that this holds for all $\delta \in \mathbb{N}$. Let $\delta \geq d_r - 1$ and assume that this identity holds for all integers up to δ. Then, the identity concluding the previous paragraph implies that

$$\begin{aligned}\sum_{0\leq k\leq d} p_k S_{\delta+1+k} &= \sum_{0\leq k\leq d} p_k \left(\sum_{0\leq j<d_r} S_{\delta+1+k-d_r+j} Q_j\right)\\ &= \sum_{0\leq j<d_r} \underbrace{\left(\sum_{0\leq k\leq d} p_k S_{\delta+1-d_r+j+k}\right)}_{=0 \text{ since } \delta+1-d_r+j\leq\delta} Q_j = 0,\end{aligned}$$

and the proof is complete. □

(A similar result is in [10, Thm. 4.5].)

The fast computation of matrix generators is usually handled via algorithms for computing minimal approximant bases [12, 10, 4]. The next result gives the main idea behind this approach. This result is similar to [10, Thm. 4.6] (see also [10, Thm. 4.7, 4.8, 4.9, 4.10]), but in some sense the reversal is on the input sequence rather than on the output matrix generator.

We recall from [11, 1] that, given a matrix $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ and an integer $d \in \mathbb{N}$, the set of *approximants for* $\mathbf{F}$ *at order* d is defined as

$$\mathcal{A}(\mathbf{F}, d) = \{\mathbf{p} \in \mathbb{K}[X]^{1\times m} \mid \mathbf{pF} = 0 \bmod X^d\}.$$

Then, the next theorem shows that relations for $\mathcal{S}$ can be retrieved as subvectors of approximants at order about $d_\ell + d_r$ for a matrix involving the first $d_\ell + d_r$ entries of $\mathcal{S}$.

Theorem 1.7. *Let $\mathcal{S} = (S_k)_k \subset \mathbb{K}^{m\times n}$ be a linearly recurrent sequence and let $(d_\ell, d_r) \in \mathbb{N}^2$ be such that the left (resp. right) canonical matrix generator of $\mathcal{S}$ has degree $\leq d_\ell$ (resp. $\leq d_r$).*

For $d > 0$, define

$$\mathbf{F} = \begin{bmatrix} \sum_{0\leq k<d} S_k X^{d-k-1} \\ -\mathbf{I}_n \end{bmatrix} \in \mathbb{K}[X]^{(m+n)\times n}. \tag{1}$$

For any relation $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ for $\mathcal{S}$, there exists $\mathbf{r} \in \mathbb{K}[X]^{1\times n}$ such that $\deg(\mathbf{r}) < \deg(\mathbf{p})$ and $[\mathbf{p}\;\;\mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$. Assuming $d \geq d_r + 1$, for any vectors $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ and $\mathbf{r} \in \mathbb{K}[X]^{1\times n}$, if $[\mathbf{p}\;\;\mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$ and $\deg([\mathbf{p}\;\;\mathbf{r}]) \leq d - d_r - 1$, then $\mathbf{p}$ is a relation for $\mathcal{S}$.

As a corollary, if $\mathbf{B} \in \mathbb{K}[X]^{(m+n)\times(m+n)}$ is a basis of $\mathcal{A}(\mathbf{F}, d_\ell + d_r + 1)$, then

- *if $\mathbf{B}$ is in Popov form then its $m \times m$ leading principal submatrix is the canonical matrix generator for $\mathcal{S}$;*
- *if $\mathbf{B}$ is row reduced then it has exactly m rows of degree $\leq d_\ell$, and the corresponding submatrix $[\mathbf{P} \;\; \mathbf{R}]$ of $\mathbf{B}$ is such that $\mathbf{P} \in \mathbb{K}[X]^{1 \times m}$ is a minimal matrix generator for $\mathcal{S}$.*

Proof. From Lemma 1.2, if $\mathbf{p}$ is a relation for $\mathcal{S}$ then $\mathbf{q} = \mathbf{pS}$ has polynomial entries, where $\mathbf{S} = \sum_{k \geq 0} S_k X^{-k-1}$. Then, the vector $\mathbf{r} = -\mathbf{p}(\sum_{k \geq d} S_k X^{d-k-1})$ has polynomial entries, has degree less than $\deg(\mathbf{p})$, and is such that $[\mathbf{p} \;\; \mathbf{r}]\mathbf{F} = \mathbf{q}X^d$, hence $[\mathbf{p} \;\; \mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$.

Conversely, if $[\mathbf{p} \;\; \mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$ we have $\mathbf{p}(\sum_{0 \leq k < d} S_k X^{d-k-1}) = \mathbf{r} \bmod X^d$. Since $d \geq d_r + 1$ and $\deg([\mathbf{p} \;\; \mathbf{r}]) \leq d - d_r - 1$, this implies that the coefficients of degree $d - d_r$ to $d - 1$ of $\mathbf{p}(\sum_{0 \leq k < d} S_k X^{d-k-1})$ are zero. Then, Lemma 1.6 shows that $\mathbf{p}$ is a relation for $\mathcal{S}$.

Finally, the two items are straightforward consequences. □

Then, using fast approximant basis algorithms, we obtain the next result.

Corollary 1.8. *Let $\mathcal{S} \subset \mathbb{K}^{m \times n}$ be a linearly recurrent sequence and let $d = d_\ell + d_r + 1$, where $(d_\ell, d_r) \in \mathbb{N}^2$ are such that the left (resp. right) canonical matrix generator of $\mathcal{S}$ has degree $\leq d_\ell$ (resp. $\leq d_r$). Then,*

- *using the algorithm of [3]: if $n \in \Omega(m)$, a left minimal matrix generator for $\mathcal{S}$ can be computed in $O(n^\omega \mathsf{M}(d) \log(d))$ operations in $\mathbb{K}$;*
- *using the algorithm of [15]: if $n \in O(m)$, a left minimal matrix generator for $\mathcal{S}$ can be computed in $O(m^\omega \mathsf{M}(nd/m) \log(nd))$ operations in $\mathbb{K}$;*
- *using the algorithm of [5]: the left canonical matrix generator for $\mathcal{S}$ can be computed in $O((m+n)^{\omega-1} \mathsf{M}(nd) \log(nd)^3)$ operations in $\mathbb{K}$.*

Note: The last cost bound comes from [5, Thm. 1.4]. But actually in the present case with the uniform shift and uniform order we may as well use a slight modification of [3, 15] (plus a call to SarSto11 to find the degrees); this is what I did in the implementation. Then, same cost as in the first item, but finds the canonical generator. Unfortunately this is not yet in the literature so this would need some details, which we want to avoid especially if we don't really care about log factors.

References

[1] B. Beckermann and G. Labahn. A uniform approach for the fast computation of matrix-type Padé approximants. *SIAM J. Matrix Anal. Appl.*, 15(3):804–823, 1994.

[2] D. S. Dummit and R. M. Foote. *Abstract Algebra.* John Wiley & Sons, 2004.

[3] P. Giorgi, C.-P. Jeannerod, and G. Villard. On the complexity of polynomial matrix computations. In *ISSAC'03*, pages 135–142. ACM, 2003.

[4] P. Giorgi and R. Lebreton. Online order basis algorithm and its impact on the block Wiedemann algorithm. In *ISSAC'14*, pages 202–209, New York, NY, USA, 2014. ACM.

[5] C.-P. Jeannerod, V. Neiger, É. Schost, and G. Villard. Fast computation of minimal interpolation bases in Popov form for arbitrary shifts. In *ISSAC'16*, pages 295–302. ACM, 2016.

[6] T. Kailath. *Linear Systems*. Prentice-Hall, 1980.

[7] E. Kaltofen and G. Villard. On the complexity of computing determinants. In *ISSAC'01*, pages 13–27. ACM, 2001.

[8] V. M. Popov. Invariant description of linear, time-invariant controllable systems. *SIAM Journal on Control*, 10(2):252–264, 1972.

[9] E. Thomé. Subquadratic computation of vector generating polynomials and improvement of the block Wiedemann algorithm. *J. Symbolic Comput.*, 33(5):757–775, 2002.

[10] W. J. Turner. *Black box linear algebra with the LINBOX library*. PhD thesis, North Carolina State University, 2002.

[11] M. Van Barel and A. Bultheel. A general module theoretic framework for vector M-Padé and matrix rational interpolation. *Numer. Algorithms*, 3:451–462, 1992.

[12] G. Villard. Further analysis of Coppersmith's block Wiedemann algorithm for the solution of sparse linear systems (extended abstract). In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation*, ISSAC '97, pages 32–39, New York, NY, USA, 1997. ACM.

[13] G. Villard. A study of Coppersmith's block Wiedemann algorithm using matrix polynomials. Technical report, LMC-IMAG, Report 975 IM, 1997.

[14] W. A. Wolovich. *Linear Multivariable Systems*, volume 11 of *Applied Mathematical Sciences*. Springer-Verlag New-York, 1974.

[15] W. Zhou and G. Labahn. Efficient algorithms for order basis computation. *J. Symbolic Comput.*, 47(7):793–819, 2012.

A Attic

For more details:

- [12, Sec. 1] when the sequence is of the form $\mathcal{S} = (\mathbf{U}\mathbf{A}^k\mathbf{V})_k$. Note that in this case the generating series can be written $\mathbf{S} = \mathbf{U}(X\mathbf{I} - \mathbf{A})^{-1}\mathbf{V}$. Link with so-called realizations from control theory [6]. . .
- [10, Chap. 4] has things related to Hankel matrices (but it is extremely detailed, including many properties which are actually about polynomial matrices and completely independent of the "linear recurrence" context)

alternative definition from [9].

Problem 1 – *Minimal matrix generator*

Input:

- sequence $\mathcal{S} = (S_k)_k \subset \mathbb{K}^{m\times n}$,
- degree bounds $(d_\ell, d_r) \in \mathbb{N}^2$.

Assumptions:

- the sequence $\mathcal{S}$ is linearly recurrent,
- the left (resp. right) canonical matrix generator of $\mathcal{S}$ has degree at most d_ℓ (resp. d_r).

Output: a minimal matrix generator for $\mathcal{S}$.

Definition A.1 ([9]). *Let $\mathcal{S} = (S_k)_{k\in\mathbb{N}} \subset \mathbb{K}^{m\times n}$ be a sequence of $m \times n$ matrices over $\mathbb{K}$. We define the generating series $\mathbf{S} = \sum_{k\geq 0} S_k X^k \in \mathbb{K}[\![X]\!]^{m\times n}$. Then, a vector $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ is said to be a* (linear recurrence) relation for $\mathcal{S}$ *if the product* $\mathbf{pS}$ *has polynomial entries, that is,* $\mathbf{pS} \in \mathbb{K}[X]^{1\times m}$.

Assume there is a nontrivial relation $\mathbf{p} = \sum_k p_k X^k$ for $\mathcal{S}$, we have

$$\sum_{k=0}^{d} p_k S_{\delta-k} = 0 \qquad \text{for all } d \geq \deg(\mathbf{p}) \text{ and } \delta \geq \max(d, \deg(\mathbf{Sp}) + 1). \tag{2}$$

The alternative definition focuses on this type of relation.

Lemma A.2. *For a given sequence $\mathcal{S} \subset \mathbb{K}^{m\times n}$, a nonzero vector $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ is a relation for Definition A.1 if and only if there exists $d \geq \deg(\mathbf{p})$ such that the reverse $X^d \mathbf{p}(X^{-1})$ is a relation for Definition 1.1.*

Proof. First, we assume that $X^d \mathbf{p}(X^{-1}) = \sum_{k=0}^{d} p_{d-k} X^k$ is a relation for Definition 1.1, for some integer $d \geq \deg(\mathbf{p})$. This means that, for all $\delta \geq 0$, we have $0 = \sum_{k=0}^{d} S_{\delta+k} p_{d-k} = \sum_{k=0}^{d} S_{\delta+d-k} p_k$. This implies that $\mathbf{Sp}$ has polynomial entries (and $\deg(\mathbf{Sp}) \leq d$).

Now, assume that $\mathbf{p}$ is a relation for Definition A.1. Taking $d = \max(\deg(\mathbf{p}), \deg(\mathbf{Sp})+1)$ in Eq. (2), we obtain $\sum_{k=0}^{d} S_{\delta-k} p_k = 0$ for all $\delta \geq d$. This implies $\sum_{k=0}^{d} S_{\delta-d+k} p_{d-k} = 0$ for all $\delta \geq d$, or equivalently, $\sum_{k=0}^{d} S_{\delta+k} p_{d-k} = 0$ for all $\delta \geq 0$. Therefore the reverse $X^d \mathbf{p}(X^{-1})$ is a relation for Definition 1.1. □