Matrix Berlekamp-Massey

November 7, 2017

1 Linearly recurrent matrix sequences

When considering sequences of matrices over a field $\mathbb{K}$, *linearly recurrent sequences* and *minimal generators* are defined similarly as for scalar sequences.

Definition 1.1 ([4, Sec. 3]). *Let $\mathcal{S} = (S_k)_{k\in\mathbb{Z}_{\geq 0}} \subset \mathbb{K}^{m\times n}$ be a matrix sequence and let $\mathbf{p} = \sum_{0\leq k\leq d} p_k X^k \in \mathbb{K}[X]^{1\times m}$ be a polynomial vector. Then,*

- $\mathbf{p}$ *is said to be a* (vector) relation for $\mathcal{S}$ *if $\sum_{k=0}^{d} p_k S_{\delta+k} = 0$ for all $\delta \geq 0$;*
- $\mathcal{S}$ *is said to be* linearly recurrent *if there exists a nontrivial relation for $\mathcal{S}$.*

Hereafter, the word generator *never* refers to these relations. This word will be reserved to generating sets for the set of all relations.

We will now see that, as in the scalar case, one can define the minimal generator of a sequence as the denominator in some irreducible fraction description of the generating series of the sequence. The set of relations for $\mathcal{S}$ is a $\mathbb{K}[X]$-submodule of $\mathbb{K}[X]^{1\times m}$, which has rank m if $\mathcal{S}$ is linearly recurrent. This is showed in [5, Fact 1] or in [8] for sequences having a *scalar* recurrence relation. Let us now link this with the property of being linearly recurrent as in the above definition.

Lemma 1.2. *The sequence $\mathcal{S}$ is linearly recurrent if and only if there exists a polynomial $P(X) = \sum_{0\leq k\leq d} p_k X^k \in \mathbb{K}[X]$ such that $\sum_{k=0}^{d} p_k S_{\delta+k} = 0$ for all $\delta \geq 0$.*

Proof. I cannot find this result in the literature..! actually, is it even true? (but we don't care since in our case the sequence obviously admits a scalar relation) □

Then, the fact that the relation module has rank m is straightforward: since the sequence is linearly recurrent, it has a scalar recurrence polynomial $P(X)$, hence for each i the vector $[0 \ \cdots \ 0 \ P(X) \ 0 \ \cdots \ 0]$ with $P(X)$ at position i is a relation for $\mathcal{S}$.

A *matrix generator* for the sequence is a matrix whose rows form a generating set for the module of relations; it is said to be

- *minimal* if the matrix is row reduced [11, 3];
- *ordered weak Popov* if the matrix is in weak Popov form [6] with pivots on the diagonal;
- *canonical* if the matrix is in Popov form [7, 3].

In this context, an important quantity related to the sequence is the determinantal degree $\deg(\det(\mathbf{P}))$, invariant for all generators $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$. Hereafter, we denote it by $\Delta(\mathcal{S})$.

Lemma 1.3. *Consider a matrix sequence $\mathcal{S} = (S_k)_{k \in \mathbb{Z}_{\geq 0}} \subset \mathbb{K}^{m \times n}$ and its generating series $\mathbf{S} = \sum_{k \geq 0} S_k / X^{k+1} \in \mathbb{K}[\![X^{-1}]\!]^{m \times n}$. Then, a vector $\mathbf{p} \in \mathbb{K}[X]^{1 \times m}$ is a 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 $\mathbf{p} = \sum_{0 \leq k \leq d} p_k X^k$. For $\delta \geq 0$, the coefficient of $\mathbf{q}$ of degree $-\delta - 1 < 0$ is $\sum_{0 \leq k \leq d} p_k S_{k+\delta}$. Hence the equivalence, by definition of a relation. The degree comparison is clear since $\mathbf{S}$ has only terms of (strictly) negative degree. □

Corollary 1.4. *A matrix sequence $\mathcal{S} = (S_k)_{k \in \mathbb{Z}_{\geq 0}} \subset \mathbb{K}^{m \times n}$ is linearly recurrent if and only if its generating series $\mathbf{S} = \sum_{k \geq 0} S_k / X^{k+1} \in \mathbb{K}[\![X^{-1}]\!]^{m \times n}$ 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 $\mathrm{rdeg}(\mathbf{Q}) < \mathrm{rdeg}(\mathbf{P})$ and $\deg(\det(\mathbf{P})) \geq \Delta(\mathcal{S})$,*
- *$\mathbf{P}$ is a matrix generator of $\mathcal{S}$ if and only if the fraction is irreducible (i.e. $\mathbf{U}\mathbf{P} + \mathbf{V}\mathbf{Q} = \mathbf{I}$ for some $\mathbf{U}$, $\mathbf{V}$),*
- *$\mathbf{P}$ is a matrix generator of $\mathcal{S}$ if and only if $\deg(\det(\mathbf{P})) = \Delta(\mathcal{S})$.*

For more details:

- [10, 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 [3]...
- [8, 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)

2 Computing minimal matrix generators

Now, we focus on the following algorithmic problem: we are given a linearly recurrent sequence and we want to find a matrix generator. If we want our algorithm to run efficiently (or simply, in finite time), we cannot access infinitely many terms of the sequence. We therefore ask for an additional input, which one often has when considering a sequence coming from some application: a bound on the degree of any minimal matrix generator. Note that all minimal generators have the same degree: that of the canonical generator. If available, a bound on the determinantal degree $\Delta(\mathcal{S})$ is sufficient; yet better bounds can be available and will yield better efficiency. We now focus on Problem 1.

We now show how the additional information of d allows us to find a matrix generator by considering only a small chunk of the sequence, rather than all its terms.

The fast computation of matrix generators is usually handled via algorithms for computing minimal approximant bases [10, 8, 2]. The next result gives the main idea behind this approach. This result is similar to [8, Thm. 4.7, 4.8, 4.9, 4.10], but in some sense the reversal

Problem 1 – *Minimal matrix generator*

Input:

- sequence $\mathcal{S} = (S_k)_k \subset \mathbb{K}^{m\times n}$,
- degree bounds $(d_\ell, d_r) \in \mathbb{Z}^2_{\geq 0}$.

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 $\mathcal{S}$.

is on the input sequence rather than on the output matrix generator (and also this section 4.2 of [8] provides many more details related to the mechanisms and output properties in the approximant basis algorithm, which we do not consider here).

We recall from [9, 1] that, given a matrix $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ and an integer $d \in \mathbb{Z}_{\geq 0}$, 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 following lemma 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 the sequence $\mathcal{S}$. Note that these bounds d_ℓ, d_r are the same as Turner's γ_1, γ_2 [8, Def. 4.6 and 4.7].

Theorem 2.1. *Let $\mathcal{S} = (S_k)_k \subset \mathbb{K}^{m\times n}$ be a linearly recurrent sequence. For $d > 0$, define*

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

Then, 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 $[\mathbf{p}\ \ \mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$.

Now, consider $(d_\ell, d_r) \in \mathbb{Z}^2_{\geq 0}$ such that the left (resp. right) canonical matrix generator of $\mathcal{S}$ has degree at most d_ℓ (resp. d_r), and define $\mathbf{F}$ for $d = d_\ell + d_r + 1$. For any vectors $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ and $\mathbf{r} \in \mathbb{K}[X]^{1\times n}$, if $[\mathbf{p}\ \ \mathbf{r}] \in \mathcal{A}(\mathbf{F}, d_\ell + d_r + 1)$ and $\deg([\mathbf{p}\ \ \mathbf{r}]) \leq d_\ell$, then $\mathbf{p}$ is a relation for $\mathcal{S}$ and $\deg(\mathbf{r}) < \deg(\mathbf{p})$.

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 (resp. ordered weak Popov) form then its $m \times m$ leading principal submatrix is the canonical (resp. ordered weak Popov) 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 1.3, 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 $[\mathbf{p}\ \ \mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$.

Then, the three items are straightforward consequences. $\square$

Corollary 2.2. *Assuming $m = \Theta(n)$, any of these matrix generators (minimal, Popov, ...) can be computed in $O(m^{\omega}\mathsf{M}(d)\log(d))$ operations in $\mathbb{K}$, where $d = \max(d_\ell, d_r)$.*

We would prefer to say that we compute the canonical form, rather than a minimal one. In theory, exactly the same asymptotic cost bound (but not yet in the literature, so this needs some short explanation; except if we do not care about logarithmic factors then this is in the literature).

With our implementation, asking for the canonical form should induce a slowdown factor of at most 2.

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] 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.

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

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

[5] Erich Kaltofen and George Yuhasz. On the matrix Berlekamp-Massey algorithm. *ACM Trans. Algorithms*, 9(4):33:1–33:24, October 2013.

[6] T. Mulders and A. Storjohann. On lattice reduction for polynomial matrices. *J. Symbolic Comput.*, 35:377–401, 2003.

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

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

[9] 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.

[10] 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.

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