On the Analytical Structure of a Vector Sequence Generated via a Linear Recursion

Avram Sidi
Computer Science Department
Technion - Israel Institute of Technology
Haifa 32000, Israel
e-mail: asidi@cs.technion.ac.il
URL: http://www.cs.technion.ac.il/~asidi

July 2018
Abstract

In this note, we discuss the nature of a vector sequence \( \{ f_n \}_{n=0}^{\infty} \in \mathbb{C}^N \) generated by a linear recursion of the form

\[
\sum_{j=0}^{m} A_j f_{k+j} = 0, \quad k = 0, 1, \ldots,
\]

where \( A_j \in \mathbb{C}^{N \times N}, \; j = 0, 1, \ldots, m, \; A_m \) is nonsingular, and \( A_0 \neq 0 \). We also discuss the nature of the function \( f(z) \) that is defined by the infinite series \( \sum_{n=0}^{\infty} f_n z^n \).
1 Introduction

In this note, we explore the analytical structure of a sequence of vectors \( \{f_n\}_{n=0}^{\infty} \in \mathbb{C}^N \) generated recursively via

\[
\sum_{j=0}^{m} A_j f_{k+j} = 0, \quad k = 0, 1, \ldots, \quad (1.1)
\]

where

\[
A_j \in \mathbb{C}^{N \times N}, \quad j = 0, 1, \ldots, m; \quad A_m \text{ nonsingular, } A_0 \neq O. \quad (1.2)
\]

Clearly, when \( A_m \) is nonsingular, (1.1) is a true forward recursion, since now

\[
f_{k+m} = -A_m^{-1} \left( \sum_{j=0}^{m-1} A_j f_{k+j} \right), \quad k = 0, 1, \ldots.
\]

Of course, \( f_0, f_1, \ldots, f_{m-1} \) must be given as initial conditions. As will become clear soon, the solution is obtained by reducing this problem to a nonlinear (or polynomial) eigenvalue problem.

In the sequel, we use uppercase italic letters to denote matrices. We denote by \( O \) the zero matrix. We also use lowercase italic letters to denote vectors. We denote by \( 0 \) the zero vector.

2 A related nonlinear eigenvalue problem

We start with the ansatz \( f_n = u \mu^n \), where \( u \in \mathbb{C}^N \) (\( u \neq 0 \) naturally) and \( \mu \in \mathbb{C} \), both to be determined. Then (1.1) becomes

\[
\sum_{j=0}^{m} \mu^{k+j} A_j u = 0, \quad k = 0, 1, \ldots \quad (2.1)
\]

Multiplying both sides by \( \mu^{-k} \), we obtain the nonlinear (or polynomial) eigenvalue problem\(^1\)

\[
\left( \sum_{j=0}^{m} \mu^j A_j \right) u = 0. \quad (2.2)
\]

Thus, \( \mu \) must be a root of the equation \( \det \left( \sum_{j=0}^{m} \mu^j A_j \right) = 0 \) since \( u \neq 0 \).

Let us look at some simple cases first.

- **The case** \( m = 1 \): We have the generalized eigenvalue problem

\[
A_0 u = -\mu A_1 u,
\]

which becomes the regular eigenvalue problem \( A_1^{-1} A_0 u = \mu u \), when \( A_1 \) is nonsingular.

---

\(^1\)For nonlinear (or polynomial) eigenvalue problems, see Björck [1, pp. 561–563], for example.
The case $m = 2$: With $x = u$ and $y = \mu u = \mu x$, we have
\[
\begin{bmatrix}
-A_0 & O & \cdots & O \\
O & I & \cdots & O \\
\vdots & \vdots & \ddots & \vdots \\
O & O & \cdots & I
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
\end{bmatrix}
= \mu
\begin{bmatrix}
A_1 & A_2 & \cdots & A_{m-1} & A_m \\
I & O & \cdots & O & O \\
O & I & \cdots & O & O \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
O & O & \cdots & I & O
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
\end{bmatrix}.
\]

The case $m = 3$: With $x = u$, $y = \mu u = \mu x$, $z = \mu^2 u = \mu y$, we have
\[
\begin{bmatrix}
-A_0 & O & O \\
O & I & O \\
O & O & I
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
= \mu
\begin{bmatrix}
A_1 & A_2 & A_3 \\
I & O & O \\
O & I & O \\
\vdots & \vdots & \vdots \\
O & O & I
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}.
\]

For arbitrary $m$, this can be generalized in a straightforward manner. Letting
\[x^{(j)} = \mu^{j-1} u, \quad j = 1, \ldots, m \Rightarrow x^{(1)} = u, \ x^{(2)} = \mu x^{(1)}, \ldots, x^{(m)} = \mu x^{(m-1)}, \quad (2.3)\]
we have the generalized eigenvalue problem
\[Vx = \mu Wx, \quad (2.4)\]
where $V, W \in \mathbb{C}^{mN \times mN}$ and $x \in \mathbb{C}^{mN}$ are given as
\[
V = \begin{bmatrix}
-A_0 & O & \cdots & O \\
O & I & \cdots & O \\
\vdots & \vdots & \ddots & \vdots \\
O & O & \cdots & I
\end{bmatrix}, \quad W = \begin{bmatrix}
A_1 & A_2 & \cdots & A_{m-1} & A_m \\
I & O & \cdots & O & O \\
O & I & \cdots & O & O \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
O & O & \cdots & I & O
\end{bmatrix}, \quad x = \begin{bmatrix}
x^{(1)} \\
x^{(2)} \\
\vdots \\
x^{(m)}
\end{bmatrix}.
\]

The generalized eigenvalue problem in (2.4) has exactly $mN$ eigenvalues, counted according to their multiplicities, provided the matrix $W$ is nonsingular, since $Vx = \mu Wx$ can be expressed as the regular eigenvalue problem $W^{-1}Vx = \mu x$ in this case. Now, $W$ is nonsingular if and only if $A_m$ is nonsingular. This can be shown as follows: When $W$ is nonsingular, the homogeneous linear system $Wx = 0$ has $x = 0$, namely, $x^{(j)} = 0, \quad j = 1, \ldots, m$, as its only solution. Now, by (2.5), we have
\[Wx = 0 \Rightarrow \sum_{j=1}^{m} A_j x^{(j)} = 0 \quad \text{and} \quad x^{(j)} = 0, \quad j = 1, \ldots, m - 1, \]
from which,
\[x^{(1)} = \cdots = x^{(m-1)} = 0 \quad \text{and} \quad A_m x^{(m)} = 0.\]

Since $x^{(1)} = \cdots = x^{(m-1)} = 0$ are also unique solutions, we need to deal with $x^{(m)}$ only. Now, $x^{(m)} = 0$ is the only solution to the homogeneous system $A_m x^{(m)} = 0$ if and only if $A_m$ is nonsingular, which we have assumed from the start.

Note that the problem $Vx = \mu Wx$ has $\mu = 0$ as an eigenvalue only when $V$ is singular, and this happens only when $A_0$ is singular, which we have allowed. (It has $\mu = \infty$ as an eigenvalue when $A_m$ is singular, which is not our case.)
3 Structure of $f_n$

Let us assume for simplicity that $W^{-1}V$ is diagonalizable. Then it has $mN$ eigenpairs $(\mu_i, x_i)$, $i = 1, \ldots, mN$. By (2.3), the vector $x_i$ has the partitioning

$$x_i = \begin{bmatrix} x_i^{(1)} \\ x_i^{(2)} \\ \vdots \\ x_i^{(m)} \end{bmatrix} = \begin{bmatrix} u_i \\ \mu_i u_i \\ \vdots \\ \mu_i^{m-1} u_i \end{bmatrix}.$$  \hspace{1cm} (3.1)

Consequently, $f_n$ can be expressed as

$$f_n = \sum_{i=1}^{mN} c_i x_i^{(1)} \mu_i^n = \sum_{i=1}^{mN} c_i u_i \mu_i^n, \quad n = 0, 1, \ldots.$$  \hspace{1cm} (3.2)

This can be verified by simply substituting (3.2) into (1.1) and invoking (2.3).

To complete our analysis, we need to show that the $mN$ unknowns $c_i$ can be determined uniquely from the $m$ initial vectors $f_0, f_1, \ldots, f_{m-1}$. Setting $n = 0, 1, \ldots, m-1$ only in (3.2), we realize that the $c_i$ satisfy the $(mN)$-dimensional linear system

$$\sum_{i=1}^{mN} c_i u_i \mu_i^n = f_n, \quad n = 0, 1, \ldots, m - 1,$$  \hspace{1cm} (3.3)

which can be expressed in matrix notation as

$$Uc = \phi,$$

where

$$U = \begin{bmatrix} u_1 & u_2 & \cdots & u_p \\ \mu_1 u_1 & \mu_2 u_2 & \cdots & \mu_p u_p \\ \vdots & \vdots & \ddots & \vdots \\ \mu_1^{m-1} u_1 & \mu_2^{m-1} u_2 & \cdots & \mu_p^{m-1} u_p \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}, \quad \phi = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{m-1} \end{bmatrix}, \quad p = mN.$$

Comparing with (3.1), we realize that the $r_i$th column of $U$ is nothing but the eigenvector $x_i$, $i = 1, \ldots, mN$. Because we have assumed that $W^{-1}V$ is diagonalizable, the $x_i$ are linearly independent. Therefore, $U$ is nonsingular, implying that $c = U^{-1} \phi$ exists and is unique.

We note that $f_n$ in (3.2) can also be expressed as

$$f_n = \sum_{i=1}^{q} \hat{u}_i \hat{\mu}_i^n, \quad n = 0, 1, \ldots, \quad \text{for some } q \leq mN \text{ and distinct } \hat{\mu}_i,$$  \hspace{1cm} (3.4)

where $\hat{\mu}_i$ are some or all of the distinct $\mu_i$ and $\hat{u}_i$ are linear combinations of the $u_i$ that correspond to the eigenvalues $\mu_i$ that are equal to $\hat{\mu}_i$. In addition, since $mN > N$ when $m > 1$, the vectors $u_i$, and hence the vectors $\hat{u}_i$, may be linearly dependent in $\mathbb{C}^N$. This is worth noting because the eigenvectors $x_i$ are linearly independent in $\mathbb{C}^{mN}$ and $u_i$ form the first $N$ components of the respective $x_i$.

Finally, in case $W^{-1}V$ is nondiagonalizable and some or all of the distinct $\mu_j$ have multiplicities greater than unity, $f_n$ is of the form

$$f_n = \sum_{i=1}^{q} p_i(n) \hat{\mu}_i^n, \quad n = 0, 1, \ldots.$$  \hspace{1cm} (3.5)

For each $i$, $p_i(n)$ is a vector-valued polynomial in $n$ of degree $\leq r_i$, where $r_i$ is the geometric multiplicity of $\hat{\mu}_i$. We leave out the details of this case.
4 Connection with vector-valued rational functions

Let us now consider the function \( f : \mathbb{C} \to \mathbb{C}^N \), defined as

\[
 f(z) = \sum_{n=0}^{\infty} f_n z^n, \quad z \in \mathbb{C},
\]

with the \( f_n \) as described in (1.1)–(1.2). Note that, due to the fact that \( f_n \) has the structure shown in (3.4) or (3.5), the series \( \sum_{n=0}^{\infty} f_n z^n \) converges absolutely and uniformly for all \( z \) sufficiently close to zero.

First, let us look at the special case in which the matrix \( W^{-1}V \) is diagonalizable. Then \( f_n \) is as in (3.2). Consequently, \( f(z) \) is a vector-valued rational function given as

\[
 f(z) = \sum_{i=1}^{mN} c_i u_i, \quad z 
\]

Whether \( W^{-1}V \) is diagonalizable or not, \( f(z) \) has an interesting structure described in the next theorem:

**Theorem 4.1** With \( f_n \) as in (1.1)–(1.2), define

\[
 s_{-1}(z) = 0; \quad s_k(z) = \sum_{n=0}^{k} f_n z^n, \quad k = 0, 1, \ldots ,
\]

Then

\[
 f(z) = \left( \sum_{j=0}^{m} A_j z^{m-j} \right)^{-1} \left( \sum_{j=0}^{m} A_j z^{m-j} s_{j-1}(z) \right), \quad \forall z \not\in \{ \mu_1^{-1}, \mu_2^{-1}, \ldots , \mu_{mN}^{-1} \}.
\]

Thus, \( f(z) \) is a vector-valued rational function with a vector-valued numerator polynomial of degree at most \( mN - 1 \) and a scalar-valued denominator polynomial of degree \( mN \).

**Proof.** Multiplying (1.1) by \( z^{k+m} \), and summing over \( k \), we first have

\[
 \sum_{k=0}^{\infty} z^{k+m} \sum_{j=0}^{m} A_j f_{k+j} = 0,
\]

which is valid for all \( z \) sufficiently close to zero. Upon rearranging, this becomes

\[
 \sum_{j=0}^{m} A_j z^{m-j} \sum_{k=0}^{\infty} f_{k+j} z^{k+j} = 0.
\]

By the fact that \( f(z) = s_{j-1}(z) + \sum_{k=0}^{\infty} f_{k+j} z^{k+j} \), it follows that

\[
 \sum_{j=0}^{m} A_j z^{m-j} [f(z) - s_{j-1}(z)] = 0,
\]

from which,

\[
 \left( \sum_{j=0}^{m} A_j z^{m-j} \right) f(z) = \sum_{j=0}^{m} A_j z^{m-j} s_{j-1}(z).
\]
Clearly, the matrix \( \sum_{j=0}^{m} A_j z^{m-j} \) is nonsingular if \( z \) is not one of the \( \mu_i^{-1} \), as we have already seen. The result in (4.3) follows.

Observe that the matrix \( \sum_{j=0}^{m} A_j z^{m-j} \) can be expressed as

\[
\sum_{j=0}^{m} A_j z^{m-j} = A_m \left( I + \sum_{j=1}^{m} B_j z^j \right); \quad B_j = A_m^{-1} A_{m-j}, \quad j = 1, \ldots, m.
\]

This implies

\[
\sum_{j=0}^{m} A_j z^{m-j} = A_m + O(z) \quad \text{as} \quad z \to 0,
\]

and hence

\[
\left( \sum_{j=0}^{m} A_j z^{m-j} \right)^{-1} = A_m^{-1} + O(z) \quad \text{as} \quad z \to 0,
\]

which, in turn, implies that \( \sum_{j=0}^{m} A_j z^{m-j} \) is invertible for all \( z \) sufficiently close to zero.

Now, the vector-valued polynomial \( \sum_{j=0}^{m} A_j z^{m-j} s_j(z) \) in (4.3) is of degree at most \( m - 1 \). The matrix \( (\sum_{j=0}^{m} A_j z^{m-j})^{-1} \) is a matrix-valued rational function with degree of numerator at most \( m(N-1) \) and degree of denominator \( mN \), as can be seen from

\[
(\sum_{j=0}^{m} A_j z^{m-j})^{-1} = \frac{\left( \sum_{j=0}^{m} A_j z^{m-j} \right)^{\text{adj}}}{\det \left( \sum_{j=0}^{m} A_j z^{m-j} \right)}.
\]

With this, the proof can now be completed.

Acknowledgement

The author wishes to thank Professor David Levin of Tel Aviv University for suggesting the problem treated in this note.

References