IMA Journal of Numerical Analysis Advance Access published July 1, 2014

IMA Journal of Numerical Analysis (2014) Page 1 of 21
doi:10.1093/imanum/dru001

On the finite rank and finite-dimensional representation of bounded semi-infinite Hankel operators

Moody T. Chu
Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205, USA
chu@math.ncsu.edu

and

Matthew M. Lin*
Department of Mathematics, National Chung Cheng University, Chia-Yi 621, Taiwan
*Corresponding author: mhlin@ccu.edu.tw

[Received on 20 September 2012; revised on 3 January 2014; accepted on 5 January 2014]

Bounded, semi-infinite Hankel matrices of finite rank over the space ℓ^2 of square-summable sequences occur frequently in classical analysis and engineering applications. The notion of finite rank often appears under different contexts and the literature is diverse. The first part of this paper reviews some elegant, classical criteria and establishes connections among the various characterizations of finite rank in terms of rational functions, recursion, matrix factorizations and sinusoidal signals. All criteria require $2d$ parameters, though with different meanings, for a matrix of rank d. The Vandermonde factorization, in particular, permits immediately a singular-value preserving, finite-dimensional representation of the original semi-infinite Hankel matrix and, hence, makes it possible to retrieve the nonzero singular values of the semi-infinite Hankel matrix. The second part of this paper proposes using the LDL* decomposition of a specially constructed sample matrix to find the unitarily equivalent finite-dimensional representation. This approach enjoys several advantages, including the ease of computation by avoiding infinite-dimensional vectors, the ability to reveal rank deficiency and the established pivoting strategy for stability. No error analysis is given, but several computational issues are discussed.

Keywords: Hankel operator; semi-infinite matrix; finite rank; Vandermonde factorization; orthogonalization; generating function; LDL* decomposition.

1. Introduction

The notion of Hankel operators has played an important role in relating many disparate parts of classical mathematics together (Adamjan *et al.*, 1968; Iohvidov, 1982; Peller, 1998, 2003), among which we mention the notable moment problems (Akhiezer, 1965). In recent years, Hankel operators have also found important applications in a wide range of disciplines outside mathematics. See the monographs (Iohvidov, 1982; Partington, 1988; Peller, 2003) for some general discussions. One such instance is the H^∞ techniques in control theory and systems theory where Hankel operators are employed to synthesize controllers that minimize sensitivity or achieve stabilization (Glover, 1984; Francis, 1987).

Hankel operators may appear in different forms. In many applications, it is common to cast the Hankel operators in matrix notation, finite or infinite,

$$H = \begin{bmatrix} h_0 & h_1 & h_2 & \dots \\ h_1 & h_2 & h_3 & \\ h_2 & h_3 & & \\ h_3 & & & \\ \vdots & & & \end{bmatrix}, \tag{1.1}$$

with respect to some preselected bases. That is, a Hankel matrix is characterized by the property that its entries depend only on the sum of the indices. We often identify H with the sequence $\{h_n\}_{n\geqslant 0}$ of its first column (and first row). When the underlying matrix is of finite dimension and of the Hankel structure, many numerical algorithms can be implemented more efficiently. For instance, it is well known that solving a linear system $H\mathbf{x} = \mathbf{b}$, where H is an $n \times n$ Hankel (or Toeplitz) matrix requires only $O(n^2)$ arithmetic operations (Heinig, 2001; Pan, 2001). Similarly, the singular value decomposition can be accomplished with $O(n^2 \log n)$ operations (Gragg & Reichel, 1989; Luk & Qiao, 2003). For semi-infinite Hankel matrices, however, numerical computation remains challenging. One example is that, despite our theoretical understanding on spectral properties of semi-infinite Hankel matrices (Widom, 1966), it requires serious effort to retrieve information analogous to the notion of finite-dimensional singular value decomposition (Kung, 1978; Young, 1983).

Indeed, a concern already arises when dealing with matrix-to-vector multiplication involving a semi-infinite matrix H. This inevitably involves infinite series. For the series to converge, we must impose the condition that H be a bounded linear operator. In most of the literature, H is considered as an operator over the space ℓ^2 of square summable (semi-infinite) sequences. It is known that a Hankel matrix H represents a bounded operator over ℓ^2 if and only if there exists a function $\psi \in L^\infty$ on the unit circle such that (Peller, 1998, Theorem 2.1)

$$h_n = \frac{1}{2\pi}\int_0^{2\pi} \psi(\theta)\, e^{-\imath n\theta}\, \mathrm{d}\theta, \quad n = 0, 1, \dots. \tag{1.2}$$

In other words, whether the sequence $\{h_n\}_{n\geqslant 0}$ for H determines a bounded operator on ℓ^2 is equivalent to whether the sequence itself represents the Fourier coefficients of an essentially bounded function over the unit disk (Nehari, 1957). Since L^∞ is dense in L^2 over the compact set $[0, 2\pi]$, we may assume without loss of generality that all rows and columns of H are sequences in ℓ^2 themselves (Bowden, 1968; Power, 1982).

In this paper, we are particularly interested in the case when a semi-infinite H is known *a priori* to have finite rank. This class of matrices arises in many fields and the literature on this subject is diverse, often under different names. A partial list of its applications includes signal processing (Golyandina *et al.*, 2001), system identification (Glover, 1984; Francis, 1987), model reduction (Kung & Lin, 1981; Markovsky *et al.*, 2005), speech and audio processing (Lemmerling & Van Huffel, 2001), modal and spectral analysis (Markovsky, 2012) and image processing (Gonzalez & Woods, 2008). Hankel matrices are also related to topics such as the Lanczos algorithm (Boley *et al.*, 1992), the recursive solution of the Yule–Walker equation (Steele, 2012) and the equivalent recursive computation of Padé approximants (Heinig & Rost, 1984; Tyrtyshnikov, 2010; Gonnet *et al.*, 2013; Ibryaeva & Adukov, 2013), and orthogonal polynomials (Akhiezer, 1965; Bultheel & Van Barel, 1997). This paper focuses only on the characterization of such a semi-infinite Hankel matrix H in terms of its nonzero singular values and singular vectors. Our main contribution is to provide a theoretical groundwork for the required computational endeavour.

We intend to bring forth two discussions. In Section 2, we review various characterizations of finite rank scattered in the literature with the aim of establishing connections in terms of rational functions, recursion, matrix factorizations and sinusoidal signals. In Section 3, we propose to form a finite-dimensional matrix that is unitarily equivalent to the original semi-infinite Hankel matrix via the rank-revealing LDL* algorithm. Numerical challenges are pointed out in the course of discussion.

2. Finite rank

As the phrase 'finite rank' plays a key role in many applications, in the first part of this paper we explain the cause and effect of such a notion. The following brief review highlights some, but not all, classical results.

2.1 *Rational functions*

We begin with perhaps the earliest theoretical consideration. By a generating function associated with a sequence $\{a_n\}_{n \geqslant 0}$ of complex numbers, we mean the formal power series defined by

$$G(z; \{a_n\}) := \sum_{n=0}^{\infty} a_n z^n. \tag{2.1}$$

It was Kronecker who first proposed an intriguing way to check the finiteness of the rank of a Hankel matrix by examining the convergence of the associated generating function. More precisely, we state the following theorem which, in fact, also asserts explicitly the corresponding rank (Peller, 1998, Theorem 4.1).

THEOREM 2.1 The Hankel matrix H has finite rank if and only if the power series (2.1) determines a rational function. In this case,

$$\operatorname{rank}(H) = \deg(zG(z; \{h_n\})), \tag{2.2}$$

where the degree of a rational function is the maximum of the degrees of its minimal constituent polynomials.

Unfortunately, determining the convergence of an infinite series usually is easier than finding the limit of its partial sums. To find the limit analytically of the generating function $G(z; \{h_n\})$ so as to calculate the rank of a given Hankel matrix as indicated above, therefore, is posed as a challenging task.

2.2 *Recursive formula*

An equivalent but more direct observation of finite rank is through the following recursive relationship among elements of H (Gantmacher, 1959, Chapter XV, Theorem 7).

THEOREM 2.2 The Hankel matrix H is of finite rank d if and only if there exist constants $\gamma_0, \ldots, \gamma_{d-1}$ such that

$$h_i = \gamma_{d-1} h_{i-1} + \gamma_{d-2} h_{i-2} + \cdots + \gamma_o h_{i-d}, \quad i = d, d+1, \ldots \tag{2.3}$$

and d is the least integer having this property.

Assuming that the semi-infinite Hankel matrix H in Theorem 2.2 is of rank d, it can be argued that the $d \times d$ leading principal submatrix $\hat{H}$, called the *trajectory matrix* of the segment $h_0, \ldots, h_{2d-2}$, is necessarily nonsingular. Rewrite finite difference equation (2.3) in the form of a linear system

$$\begin{bmatrix} h_0 & h_1 & h_2 & \dots & h_{d-1} \\ h_1 & h_2 & h_3 & & h_d \\ h_2 & h_3 & & & h_{d+1} \\ \vdots & & & & \\ h_{d-1} & h_d & & & h_{2d-2} \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_{d-1} \end{bmatrix} = \begin{bmatrix} h_d \\ h_{d+1} \\ \vdots \\ h_{2d-1} \end{bmatrix}, \tag{2.4}$$

known as the Yule–Walker equations in signal processing, for the coefficients $\gamma_0, \gamma_1, \dots, \gamma_{d-1}$. It follows that the parameters $\gamma_0, \dots, \gamma_{d-1}$ and, consequently, the entire semi-infinite time series, are uniquely determined by the first $2d$ elements $h_0, h_1, \dots, h_{2d-1}$. In this sense, we may say that a semi-infinite Hankel matrix of finite rank d has a $2d$-dimensional representation $\{h_0, h_1, \dots, h_{2d-1}\}$. The Yule–Walker equations are ubiquitous in science, finance and technology. It is said that 'every cell phone solves the Yule–Walker equations every 10 microseconds' (Dutoit, 2004) and, needless to say, many discussions and effective algorithms are available. We mention an interesting website (Steele, 2012) where further references can be found.

In theory, for a given Hankel matrix H, we could repeatedly try out a sequence of nested linear systems in the form of (2.4) by gradually *decreasing* the value of d from a large initial guess until the trajectory matrix becomes nonsingular for the first time, whence the linear system is solved and the rank is determined. Nonetheless, issues of numerical rank do arise in reality. Even though the Courant–Fischer theorem guarantees the interlacing of singular values between two nested trajectory matrices, it does not guarantee how far a nonsingular matrix is away from the nearest singular matrix. Even without the presence of noise, a nearly singular matrix may have to be treated as singular. Rank revealing algorithms developed thus far in the literature cannot resolve this problem perfectly (Lee *et al.*, 2009; Foster & Liu, 2012).

2.3 *Sinusoidal signals*

In the field of signal processing, there is another popular way to parameterize Hankel matrices of finite rank by means of sinusoidal signals. The forward connection of associating a damped signal of finite component with a Hankel matrix of finite rank is quite straightforward. Consider a noiseless time-domain signal comprising d components of exponentially decaying sinusoids,

$$\mathfrak{s}(t) = \sum_{\ell=1}^{d} a_\ell \, e^{-\zeta_\ell t} \, e^{\imath(2\pi\nu_\ell t + \phi_\ell)}, \tag{2.5}$$

where a_ℓ, ζ_ℓ, ν_ℓ and ϕ_ℓ are real numbers denoting the magnitude, the decay rate, the frequency and the phase angle of the ℓth sinusoid, respectively. Starting with $t_0 = 0$ and sampling this signal at uniformly spaced nodes $t_0, t_1, \dots$ with fixed interval length Δt (so $1/\Delta t$ is the so-called sampling rate), we obtain an infinite sequence

$$s_k := \mathfrak{s}(t_k) = \sum_{\ell=1}^{d} a_\ell \, e^{-\zeta_\ell k \Delta t} \, e^{\imath(2\pi\nu_\ell k\Delta t + \phi_\ell)} = \sum_{\ell=1}^{d} a_\ell \, e^{\imath\phi_\ell} \left(e^{(-\zeta_\ell + \imath 2\pi\nu_\ell)\Delta t}\right)^k.$$

For simplicity, denote

$$\beta_\ell := a_\ell\, e^{\iota\phi_\ell}, \tag{2.6}$$

$$\lambda_\ell := e^{(-\zeta_\ell+\iota 2\pi\nu_\ell)\Delta t}. \tag{2.7}$$

Obviously, the time series $\{s_0, s_1, \ldots\}$ enjoys the relationship

$$s_k = \sum_{\ell=1}^{d} \beta_\ell \lambda_\ell^k. \tag{2.8}$$

Such a relationship implies a decomposition of the corresponding trajectory matrix

$$S := \begin{bmatrix} s_0 & s_1 & s_2 & \cdots \\ s_1 & s_2 & & \\ s_2 & & & \\ \vdots & & & \end{bmatrix} = \sum_{\ell=1}^{d} \beta_\ell \begin{bmatrix} \lambda_\ell^0 & \lambda_\ell^1 & \lambda_\ell^2 & \ldots \\ \lambda_\ell^1 & \lambda_\ell^2 & & \\ \lambda_\ell^2 & & & \\ \vdots & & & \end{bmatrix}$$

into the form

$$S = \begin{bmatrix} \lambda_1^0 & \lambda_2^0 & \ldots & \lambda_d^0 \\ \lambda_1^1 & \lambda_2^1 & & \lambda_d^1 \\ \lambda_1^2 & & & \\ \vdots & & & \vdots \end{bmatrix} \begin{bmatrix} \beta_1 & 0 & \ldots & 0 \\ 0 & \beta_2 & & \\ \vdots & & \ddots & \\ 0 & & & \beta_d \end{bmatrix} \begin{bmatrix} \lambda_1^0 & \lambda_1^1 & \lambda_1^2 & \ldots \\ \lambda_2^0 & \lambda_2^1 & & \\ \vdots & & & \\ \lambda_d^0 & \lambda_d^1 & & \ldots \end{bmatrix}, \tag{2.9}$$

which is known as the Vandermonde factorization of S. Under the generic assumption that all $\lambda_1, \ldots, \lambda_d$ are distinct, it is seen immediately from (2.9) that the Hankel matrix S is necessarily of rank d. Note also that for strictly decaying signals, we should have $\zeta_\ell > 0$ and, hence, $|\lambda_\ell| < 1$, implying that S is a bounded operator.

It turns out that the backward connection of associating a Hankel matrix of finite rank to a damped signal of finite component also holds. The reason is based on the fact that any given Hankel matrix of finite rank d enjoys a Vandermonde factorization of which (2.9) is a special case. We shall establish the details in Section 2.4. When this factorization is in hand, then by (2.6) the polar form of the complex number β_ℓ determines the magnitude a_ℓ and phase angle ϕ_ℓ. In the meantime, by (2.7) the polar form of the quantity λ_ℓ determines, up to a scaling by Δt, the decay rate ζ_ℓ and the frequency ν_ℓ. That there is a dependence on Δt makes sense because the measurement of decay rate and frequency should be relative to the meaning of a unit time that must be defined somewhere. Thus, once Δt is specified, a composite signal is completely determined from a given Vandermonde decomposition. In short, we can go back and forth interchangeably between a bounded low rank Hankel matrix and a sinusoidal signal through the relationship (2.8). Ultimately, the term 'singular values' of a signal $\mathfrak{s}$ can be understood from those of the corresponding trajectory matrix S.

The role of the sampling rate is significant and is worth additional discussion. Out of the very same composite signal $\mathfrak{s}(t)$, a large Δt, that is, a low sampling rate, will make the numbers λ_ℓ cluster around the origin, making the sequence $\{s_k\}$ quickly decay to zero. The corresponding Hankel matrix S would then contain many anti-diagonals whose entries are nearly machine zero. In theory, as we have described above, the Hankel matrix S should still be of rank d. When we calculate its numerical rank, however, the number of calculated singular values that are greater than a specified threshold could be far less than d (see Example 2 in Section 4). Such a mismatch then causes confusion on how the original signal should be represented.

More difficulties arise when noise is present in a signal. Even with a slight perturbation, the above connection usually is lost, that is, the trajectory matrix of the contaminated time series usually is of full rank and the true value of d is smudged. Reconstructing a Hankel low rank approximation to the contaminated trajectory matrix is a problem of practical importance and has been extensively studied in the literature. Far from being complete, we mention only a few references (Park *et al.*, 1999; Lemmerling & Van Huffel, 2001; Chu *et al.*, 2003; Markovsky *et al.*, 2005; Gillard & Zhigljavsky, 2011; Markovsky, 2012). See also (Auvergne, 1988; Vautard *et al.*, 1992; Allen & Smith, 1996; Mineva & Popivanov, 1996; Danilov, 1997; Ghil & Taricco, 1997; Yiou *et al.*, 2000) for the popular singular-value spectrum analysis (SSA) techniques. In all, it is imperative to first determine an appropriate rank before any filtering technique can be applied to get rid of the noise. Such a task of rank detection remains a challenging question to this date (Golub & Van Loan, 2013).

2.4 *Vandermonde factorization*

In this section, we fill in the details that associate a semi-infinite Hankel matrix of finite rank with a sinusoidal signal. We approach this backward connection by exploiting finite difference equation (2.3) and the corresponding characteristic polynomial

$$p(\lambda) := \lambda^d - \gamma_{d-1}\lambda^{d-1} - \cdots - \gamma_1\lambda - \gamma_0. \tag{2.10}$$

To fix the ideas, let λ_ℓ, $\ell = 1, \ldots, r$, denote the distinct roots of $p(\lambda)$, each of which has multiplicity ρ_ℓ. So $\sum_{\ell=1}^{r} \rho_\ell = d$. A general solution to difference equation (2.3) can be formulated via the superposition principle as follows. Let $[\beta_1^{(0)}, \ldots, \beta_1^{(\rho_1-1)}, \beta_2^{(0)}, \ldots, \beta_2^{(\rho_2-1)}, \ldots \beta_r^{(0)}, \ldots \beta_r^{(\rho_r-1)}]^{\mathrm{T}} \in \mathbb{C}^d$ be the solution to the confluent Vandermonde system (Kalman, 1984)

$$\mathscr{V}_{p(\lambda)} \begin{bmatrix} \beta_1^{(0)} \\ \vdots \\ \beta_1^{(\rho_1-1)} \\ \vdots \\ \beta_r^{(0)} \\ \vdots \\ \beta_r^{(\rho_r-1)} \end{bmatrix} = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{d-1} \end{bmatrix}, \tag{2.11}$$

where $\mathscr{V}_{p(\lambda)} \in \mathbb{C}^{d \times d}$ is composed of r blocks of submatrices

$$\mathscr{V}_{p(\lambda)} = [\mathscr{V}^{(1)}, \ldots, \mathscr{V}^{(r)}]$$

with the block $\mathscr{V}^{(\ell)} \in \mathbb{C}^{d\times\rho_\ell}$, $\ell = 1, \ldots, r$, defined by

$$\mathscr{V}^{(\ell)} := \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ \lambda_\ell & 1 & 0 & \ldots & \\ \lambda_\ell^2 & 2\lambda_\ell & 1 & \ldots & \\ \lambda_\ell^3 & 3\lambda_\ell^2 & 3\lambda_\ell & & \\ \vdots & & & \ddots & \\ \lambda_\ell^{\rho_\ell-1} & & & & 1 \\ \lambda_\ell^{\rho_\ell} & & & & \rho_\ell\lambda_\ell \\ \vdots & \vdots & & & \vdots \\ \lambda_\ell^{d-1} & (d-1)\lambda_\ell^{d-2} & \dfrac{(d-1)(d-2)}{2}\lambda_\ell^{d-3} & \ldots & \begin{pmatrix} d-1 \\ \rho_\ell - 1 \end{pmatrix}\lambda_\ell^{d-\rho_\ell} \end{bmatrix}.$$

For notational simplicity, entries of the matrix $\mathscr{V}^{(\ell)} = [v_{ij}^{(\ell)}]$ can be expressed as

$$v_{ij}^{(\ell)} = c_{ij}\lambda_\ell^{i-j}, \quad i = 0, 1, \ldots, d-1,\ j = 0, 1, \ldots, \rho_\ell - 1, \tag{2.12}$$

where $c_{ij} = 0$ if $i < j$; and $c_{ij} = i!/(i-j)!j!$ otherwise. Note also that for convenience, the indices of the entries start from $(0, 0)$. The purpose of imposing system (2.11) is to enforce the initial values $h_0, \ldots, h_{d-1}$ in the solution to finite difference equation (2.3). Trivially, the equations in (2.11) are equivalent to

$$h_i = \sum_{\ell=1}^{r}\sum_{j=0}^{\rho_\ell-1}\beta_\ell^{(j)}c_{ij}\lambda_\ell^{i-j}, \quad i = 0, 1, \ldots, d-1,$$

from which we can verify recursively that

$$\begin{aligned} h_i &= \sum_{s=0}^{d-1}\gamma_s h_{i-d+s} \\ &= \sum_{s=0}^{d-1}\gamma_s\left(\sum_{\ell=1}^{r}\sum_{j=0}^{\rho_\ell-1}\beta_\ell^{(j)}c_{i-d+s,j}\lambda_\ell^{i-d+s-j}\right) = \sum_{\ell=1}^{r}\sum_{j=0}^{\rho_\ell-1}\beta_\ell^{(j)}\left(\sum_{s=0}^{d-1}\left(c_{i-d+s,j}\gamma_s\lambda_\ell^{i-d+s-j}\right)\right) \\ &= \sum_{\ell=1}^{r}\sum_{j=0}^{\rho_\ell-1}\beta_\ell^{(j)}c_{ij}\lambda_\ell^{i-j}, \quad i = d, d+1, \ldots, \end{aligned} \tag{2.13}$$

where the last equality is obtained by using the identity

$$\frac{d^j(\lambda_\ell^{i-d}p(\lambda_\ell))}{d\lambda^j} = 0, \quad \ell = 1, \ldots, r,\ j = 0, 1, \ldots, \rho_\ell - 1 \tag{2.14}$$

for all $i \geqslant d$. Let V_∞ denote the confluent Vandermonde matrix of size $\infty \times d$ obtained by extending the matrix $\mathscr{V}_{p(\lambda)}$ defined in (2.12) downward to infinite length. Then the following result, known as the Vandermonde factorization for a semi-infinite Hankel matrix H, is but a rearrangement of expression (2.13) in a matrix form (Boley *et al.*, 1997).

THEOREM 2.3 Suppose H is a semi-infinite Hankel matrix of rank d. Then there exists a $d \times d$ block diagonal matrix D_∞ whose ℓth block is of size $\rho_\ell \times \rho_\ell$ and is Hankel and upper anti-triangular such that

$$H = V_\infty D_\infty V_\infty^{\mathrm{T}}. \tag{2.15}$$

It is important to point out that the Vandermonde factorization described above is valid for any semi-infinite Hankel matrix of rank d. If H is further required to be bounded (as an operator), then the root conditions must be satisfied, that is, all roots λ_ℓ of the polynomial (2.10) must have modulus less than or equal to one and those of modulus one must be simple.

Consider the generic case, as we shall assume henceforth, in which all roots of $p(\lambda)$ are distinct. The general solution (2.13) can be easily reduced to

$$\begin{aligned} h_i &= \sum_{s=0}^{d-1} \gamma_s h_{i-d+s} = \sum_{s=0}^{d-1} \gamma_s \left(\sum_{\ell=1}^{d} \beta_\ell^{(0)} \lambda_\ell^{i-d+s} \right) = \sum_{\ell=1}^{d} \beta_\ell^{(0)} \left(\sum_{s=0}^{d-1} \gamma_s \lambda_\ell^{i-d+s} \right) \\ &= \sum_{\ell=1}^{d} \beta_\ell^{(0)} \lambda_\ell^{i-d} \left(\sum_{s=0}^{d-1} \gamma_s \lambda_\ell^{s} \right) = \sum_{\ell=1}^{d} \beta_\ell^{(0)} \lambda_\ell^{i}, \quad i = d, d+1, \ldots, \end{aligned} \tag{2.16}$$

which now formally agrees with (2.8). In this way, we may now associate the Hankel matrix with a corresponding sinusoidal signal.

The ideas of truncation and the extension are worth mentioning. Given a semi-infinite Hankel matrix H with rank d, any of its $m \times k$ submatrix with $\min\{m, k\} \geqslant d$ is of rank at most d. However, not all low rank finite Hankel matrices can be embedded as a submatrix of a semi-infinite Hankel matrix with equal rank. It seems intuitive that one could trivially expand a given $m \times k$ Hankel matrix into a semi-infinite Hankel matrix by padding with innocuous zeros along the anti-diagonals, but such an extension usually results in a semi-infinite matrix with higher rank. Such a study actually is in the area of singular extension theory (Iohvidov, 1982, II.9) and has applications such as Gaussian quadrature (Golub & Welsch, 1969). In the course of our discussion, we have already suggested an alternative method of extension by embedding a given $d \times d$ nonsingular Hankel matrix as the leading principal submatrix of a nontrivial semi-infinite Hankel matrix H with a specified value for one additional element h_{2d-1}. After solving corresponding equation (2.4), we will have created the recursive relationship (2.13) which results in a matrix H of rank d. In this way, any finite-dimensional nonsingular Hankel matrix also admits a Vandermonde factorization (Boley *et al.*, 1997).

Finally, we mention in passing that our discussion thus far has been about scalar Hankel matrices. Similar representation can be extended to nonsingular block Hankel matrices (Kung & Lin, 1981; Feldmann & Heinig, 1996). For the purpose of conveying our basic ideas, we shall concentrate on the scalar Hankel matrices only in this paper. The generalization to block Hankel matrices deserves further investigation.

3. Finite-dimensional representation

Given a bounded Hankel operator H of finite rank d, we see immediately from the Vandermonde factorization (2.15) that

$$\mathfrak{R}(H) = \mathfrak{R}(V_\infty), \tag{3.1}$$

$$\mathfrak{R}(H^*) = \mathfrak{R}(\bar{V}_\infty), \tag{3.2}$$

$$\begin{tikzcd}
\mathfrak{R}(\overline{V}_\infty) \arrow[r, "H|_{\mathfrak{R}(\overline{V}_\infty)}"] \arrow[d, hook, "id"'] & \mathfrak{R}(V_\infty) \arrow[d, hook, "id"] \\
\ell^2 \arrow[r, "H"] & \ell^2
\end{tikzcd}$$

FIG. 1. Commutative diagram relating the semi-infinite Hankel operator H of finite rank and its corresponding reduced map $H|_{\mathfrak{R}(\bar{V}_\infty)}$.

where $\bar{V}_\infty$ denotes the complex conjugate of V_∞ and $\mathfrak{R}(H)$ stands for the range space of H. By the Fredholm alternative theorem, we see that H and H^* have effects only on the finite-dimensional spaces $\mathfrak{R}(\bar{V}_\infty)$ and $\mathfrak{R}(V_\infty)$, respectively. Thus, it is natural to expect that H should have a finite-dimensional representation. We stress that we are interested in rendering the 'action' H as a map, which is more than just a passive characterization of H by finitely many parameters, such as the first $2d$ elements $\{h_0, h_1, \ldots, h_{2d-1}\}$ or the corresponding Vandermonde parameters $\{\beta_1, \ldots, \beta_d, \lambda_1, \ldots, \lambda_d\}$. In particular, we are interested in finding a singular value decomposition of H and representing H via a finite-dimensional matrix of size $d \times d$.

For the singular value computation, an interesting approach employing the generating function (2.1) of the infinite sequence $\{h_0, h_1, \ldots\}$ in ℓ^2 has already been proposed in Young (1983). That development was based on the fact asserted by Theorem 2.1 that the associated generating function of H is necessarily a rational function, whence the range space $R(H)$ of H as well as that of its adjoint H^* can be fully characterized in terms of some specially selected finite bases of rational functions. Our approach in this note is similar in spirit, but different from Young (1983), in that we utilize the Vandermonde factorization (2.15) of H to build the representation.

The connections described in the preceding section can be implemented as a numerical procedure to achieve the Vandermonde factorization of a given Hankel matrix H. We caution, however, that many computational issues remain to be thoroughly investigated, including the stability of the algorithm itself that leads to the Vandermonde factorization and the conditioning of the roots of the characteristic polynomial (2.10). These are important and legitimate questions that, to our knowledge, have only been partially studied in the literature (Kibangou & Favier, 2007; Bezerra, 2012). In order to stay focused on the singular value decomposition, we shall not discuss these issues in this note either, but assume that the factorization is already in hand. We hope that the following discussion serves at least as a theoretic framework of interest.

To characterize the action of H, it suffices to find a matrix representation of the restricted map

$$H|_{\mathfrak{R}(\bar{V}_\infty)} : \mathfrak{R}(\bar{V}_\infty) \to \mathfrak{R}(V_\infty)$$

with respect to some orthonormal bases for the two range spaces. The reason is that, by the Fredholm alternative theorem, the action of H on the orthogonal complement of $\mathfrak{R}(\bar{V}_\infty)$ is zero. We summarize the relationship through the commutative diagram in Fig. 1. It is important to point out that while the images of the restricted map $H|_{\mathfrak{R}(\bar{V}_\infty)}$ are identical to those of H as an operator on nontrivial elements, it no longer has the Hankel structure in general.

Since the restricted map is over finite-dimensional subspaces, we now describe how its finite-dimensional matrix representation $\tilde{H}$ can be found.

LEMMA 3.1 Suppose that the matrix V_∞ has a QR decomposition

$$V_\infty = QR, \tag{3.3}$$

where the columns of $Q \in \mathbb{C}^{\infty\times d}$ are mutually orthonormal and $R \in \mathbb{C}^{d\times d}$ is upper triangular. Then the nonzero singular values of H are the same as those of the $d \times d$ matrix

$$\tilde{H} := Q^* H \bar{Q} = RD_\infty R^{\mathrm{T}}. \tag{3.4}$$

Proof. The two subspaces $\mathfrak{R}(V_\infty)$ and $\mathfrak{R}(\bar{V}_\infty)$ have columns of Q and $\bar{Q}$ as orthonormal bases, respectively. The first equality follows from the standard formula for change of basis, i.e., $H\bar{Q} = Q\tilde{H}$. The second equality follows from (2.15) to (3.3), which yield $H = \mathrm{QRD}_\infty R^{\mathrm{T}} Q^{\mathrm{T}}$ and from the fact that $Q^* Q = Q^{\mathrm{T}} \bar{Q} = I$. □

Both equalities in (3.4) have important meanings. The first equality ensures that $\tilde{H}$ and H are unitarily equivalent and, hence, the singular values of both systems are preserved. That is, the nonzero singular values of the semi-infinite matrix H can now be calculated from the finite matrix $\tilde{H}$. To generate $\tilde{H}$ by using Q, however, is cumbersome because the multiplication involves infinite series formed from the columns of Q and H. For computation, the second equality becomes handy because both matrices R and D_∞ are of size $d \times d$ only.

We now explain how an orthonormal basis can be found for V_∞. Our ultimate goal is to compute the finite-dimensional matrix R without explicitly forming Q so that together with D_∞ we can compute the nonzero singular values of H via $\tilde{H}$. We divide our approach into three steps.

3.1 *Modified Gram–Schmidt orthogonalization*

The challenge in computing the QR decomposition for V_∞ is that we have to deal with infinite-dimensional vectors. We first propose using the Gram–Schmidt process to carry out the task. One attractive feature of this approach is that, with the Vandermonde structure, applying the Gram–Schmidt orthogonalization process to columns of V_∞ is equivalent to applying the process to the corresponding generating functions over ℓ^2. All inner products of infinite-dimensional vectors can be handled effectively via rational arithmetic operations on generating functions.

We point out two useful facts before we proceed. Firstly, observe that the generating function f_ℓ corresponding to the ℓth column of V_∞, $\ell = 1, \ldots, d$, is given by

$$f_\ell(z) := G(z; \{\lambda_\ell^n\}) = \sum_{n=0}^{\infty} \lambda_\ell^n z^n = \frac{1}{1-\lambda_\ell z}. \tag{3.5}$$

Secondly, a natural way to define an inner product between two generating functions is by

$$\langle G(z;\{a_n\}), G(z;\{b_n\}\rangle := \sum_{n=0}^{\infty} a_n \bar{b}_n, \tag{3.6}$$

where $\bar{b}_n$ denotes the complex conjugate of b_n. The relationship

$$\langle G(z;\{s^n\}), G(z;\{b_n\})\rangle = \left\langle \frac{1}{1-sz}, G(z;\{b_n\}) \right\rangle = \sum_{n=0}^{\infty} s^n \bar{b}_n = G(s, \{\bar{b}_n\}) = \overline{G(\bar{s}, \{b_n\})} \tag{3.7}$$

Algorithm 1 Modified Gram–Schmidt process on the generating functions.

Require: complex numbers $\lambda_1, \dots, \lambda_d$ {from the Vandermonde decomposition of a Hankel matrix H of rank d}

Ensure: QR decomposition of V_∞ {in the form (3.9)}

for $\ell = 1, \cdots, d$ **do**
  $w_\ell := f_\ell$
end for
for $\ell = 1, \cdots, d$ **do**
  $r_{\ell\ell} := \|w_\ell\|$ {or simply evaluate $r_{\ell\ell} := \sqrt{\overline{w_\ell(\bar{\lambda}_\ell)}}$}
  $w_\ell := \frac{w_\ell}{r_{\ell\ell}}$
  for $k = \ell + 1, \cdots, d$ **do**
    $r_{\ell k} := \langle w_k, w_\ell \rangle$ {or simply evaluate $r_{\ell k} := \overline{w_\ell(\bar{\lambda}_k)}$}
    $w_k := w_k - r_{\ell k} w_\ell$
  end for
end for

then holds whenever the evaluation $G(\bar{s}, \{b_n\})$ makes sense.

Summarized in Algorithm 1 is the usual modified Gram–Schmidt scheme. There is nothing new at first glance. However, what is the important point is that the needed inner products and norms can all be computed in terms of the generating functions. Because the infinite vectors w_ℓ are represented by functions of the variable z, most calculations can be accomplished by function evaluations. We illustrate a few steps below.

It is obvious by (3.7) that

$$r_{11} = \sqrt{\langle G(z; \{\lambda_1^n\}), G(z; \{\lambda_1^n\}) \rangle} = \sqrt{f_1(\bar{\lambda}_1)} = \frac{1}{\sqrt{1 - \lambda_1 \bar{\lambda}_1}},$$

which is guaranteed to be a positive number. After obtaining the normalized $w_1(z)$, we see by (3.7) again that the first row of R is attainable via

$$r_{1k} = \langle w_k, w_1 \rangle = \overline{w_1(\bar{\lambda}_k)}, \quad k = 2, \dots, d.$$

According to the algorithm, we temporarily should have

$$w_k(z) = f_k(z) - r_{1k} w_1(z), \quad k = 2, \dots, d,$$

each of which is orthogonal to $w_1(z)$ and, hence,

$$r_{22} = \|w_2\| = \sqrt{\langle w_2, w_2 \rangle} = \sqrt{\langle f_2 - r_{12} w_1, w_2 \rangle} = \sqrt{\overline{w_2(\bar{\lambda}_2)}}.$$

The conjugation in the above actually is not needed because in exact arithmetic the quantity $w_2(\bar{\lambda}_2)$ is guaranteed to be real and non-negative. We keep the conjugation for consistency with the general expression that, after normalizing $w_2(z)$, we can generate the second row of R via

$$r_{2k} = \langle w_k, w_2 \rangle = \langle f_k - r_{1k} w_1, w_2 \rangle = \langle f_k, w_2 \rangle = \overline{w_2(\bar{\lambda}_k)}, \quad k = 2, \ldots, d.$$

By recursion, we are ready to generate the ℓth row of R via the expression

$$r_{\ell k} = \langle w_k, w_\ell \rangle = \overline{w_\ell(\bar{\lambda}_k)}, \quad k = \ell, \ldots, d, \tag{3.8}$$

where $w_\ell(z)$ is the normalized generating function as specified in the algorithm. In this way, the QR factorization for the semi-infinite matrix V_∞ can be represented in terms of a decomposition of functions $f_1, \ldots, f_d$ in the form

$$[f_1, f_2, \ldots, f_d] = [w_1, w_2, \ldots, w_d] \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & \ldots & r_{1d} \\ 0 & r_{22} & r_{23} & r_{24} & & r_{2d} \\ 0 & 0 & r_{33} & r_{34} & & \\ \vdots & & & & & \\ & & & & \ddots & \\ 0 & & & & \ldots & r_{dd} \end{bmatrix}, \tag{3.9}$$

whereas $w_1, \ldots, w_d$ themselves are mutually orthonormal functions (infinite sequences) with respect to the inner product (3.6) and the $d \times d$ upper triangular matrix $R = [r_{\ell k}]$ is really the one needed in the finite-dimensional representation (3.4).

It is interesting to make the following observation about the roots of the function $w_\ell(z)$.

LEMMA 3.2 Let w_ℓ denote the generating function corresponding to the ℓth column of the orthogonal matrix in the QR decomposition of V_∞. Then the function $w_\ell(z)$ has at least $\ell - 1$ roots at $\bar{\lambda}_1, \ldots, \bar{\lambda}_{\ell-1}$.

Proof. In the same way as (3.8), for each $j = 1, \ldots, \ell - 1$ we can calculate the inner product $\langle w_j, w_\ell \rangle = \overline{w_\ell(\bar{\lambda}_j)}$ as simply the conjugation of the function evaluation at $\bar{\lambda}_j$. By construction, $w_\ell(z)$ is orthogonal to all w_j. □

The implication of Lemma 3.2 is that if $\bar{\lambda}_\ell$ is near to any of these roots, then the value of $r_{\ell\ell}$ will be small, causing numerical instability even for the modified Gram–Schmidt process. We shall offer an alternative way to circumvent this difficulty.

3.2 *Implicit calculation*

In Algorithm 1, we compute each function $w_\ell(z)$ in its entirety. Such a function is needed only if the corresponding column in Q is needed. Symbolic calculation of the rational function $w_\ell(z)$ is theoretically possible, but lengthy and tedious.

On the other hand, a close examination of Algorithm 1 indicates that all we need is the evaluation of $w_\ell(z)$ at discrete points $\bar{\lambda}_\ell, \ldots, \bar{\lambda}_d$ for each $\ell = 1, \ldots, d$. These are the entries for the upper triangular

Algorithm 2 Reorganized Modified Gram–Schmidt Process (Implicit Q).

Require: complex numbers $\lambda_1, \ldots, \lambda_d$ {from the Vandermonde decomposition of a Hankel matrix H of rank d}

Ensure: Hermitian matrix F, lower triangular matrix Ω, and upper triangular matrix R {all finite-dimensional and satisfying (3.12)}

$\Lambda := [\lambda_1, \ldots, \lambda_d]$
$F := 1./(1 - \Lambda' * \Lambda)$ {sample matrix}
for $\ell = 1, \cdots, d$ **do**
  $r(\ell,\ell) := \sqrt{F(\ell,\ell) - \omega(\ell, 1:\ell-1) * r(1:\ell-1,\ell)}$
  $\omega(\ell:d,\ell) := (F(\ell:d,\ell) - \omega(\ell:d, 1:\ell-1) * r(1:\ell-1,\ell))/r(\ell,\ell)$
  $r(\ell, \ell:d) := (\omega(\ell:d,\ell))'$
end for

matrix R. Clearly, we have

$$w_1(\bar{\lambda}_k) = \frac{f_1(\bar{\lambda}_k)}{\sqrt{f_1(\bar{\lambda}_1)}}, \quad k = 1, \ldots, d, \tag{3.10}$$

to begin with. The remaining function evaluations can be written recursively as

$$w_\ell(\bar{\lambda}_k) = \frac{f_\ell(\bar{\lambda}_k) - \sum_{j=1}^{\ell-1} r_{j\ell} w_j(\bar{\lambda}_k)}{\sqrt{f_\ell(\bar{\lambda}_\ell) - \sum_{j=1}^{\ell-1} r_{j\ell} w_j(\bar{\lambda}_\ell)}}, \quad \ell = 2, \ldots, d,\ k = \ell, \ldots, d. \tag{3.11}$$

Reorganized in Algorithm 2 is the very same Gram–Schmidt process, except that the generating functions $w_\ell(z)$ and, thus, the corresponding orthogonal matrix Q are now implicit. We write the code in the MATLAB syntax for the convenience of vector processing. The algorithm generates internally one Hermitian matrix F, referred to as the sample matrix, and one lower triangular matrix Ω. The (k,ℓ) entries of F and Ω are $f_\ell(\bar{\lambda}_k)$ and $w_\ell(\bar{\lambda}_k)$, denoting the evaluation of the generating functions $f_\ell(z)$ and $w_\ell(z)$ at $\bar{\lambda}_1, \ldots, \bar{\lambda}_d$ (and recall Lemma 3.2), respectively. The upper triangular matrix $R := \Omega^*$ returned by the algorithm is precisely the matrix needed in (3.9). By construction, the three matrices satisfy the relationship

$$F = \Omega R. \tag{3.12}$$

The expressions (3.9) and (3.12) look similar, but there is an important difference. In (3.9), columns of the matrices $[f_1, \ldots f_d]$ and $[w_1, \ldots, w_d]$ are supposedly infinite sequences or, in our representation, generating functions. In (3.12), columns of the matrices F and Ω are d-dimensional vectors, representing the evaluations of respective functions at finitely many points $\bar{\lambda}_1, \ldots, \bar{\lambda}_d$. Also, columns of the matrix $[w_1, \ldots, w_d]$ in (3.9) are mutually orthogonal, but it is not the case for columns of the matrix Ω.

Algorithm 2 is modified from Algorithm 1 with the desire to avoid the computation of Q. But now it can be realized as the following trivial fact which, nonetheless, fundamentally changes the tactics of computation for the finite representation (3.4), as we shall describe in Section 3.3.

THEOREM 3.3 The sample matrix F is Hermitian and positive definite and relationship (3.12) is the unique Cholesky decomposition of F with positive diagonal entries in R.

Algorithm 3 *LDL** approach.

Require: complex numbers $\lambda_1, \dots, \lambda_d$ {from the Vandermonde decomposition of Hankel matrix H of rank d}
Ensure: Hermitian matrix F, unit lower triangular matrix L, diagonal matrix D, and permutation vector P

$\Lambda := [\lambda_1, \dots, \lambda_d]$
$F := 1./(1 - \Lambda' * \Lambda)$
[L,D,P] = ldl(F,'vector')
{The following has not been implemented in this work}
if D has extremely small entries **then**
  Reduce d
  Recalculate $\lambda_1, \dots, \lambda_d$
end if

Proof. Given the Vandermonde matrix V_∞, it is easy to see that $F = V_\infty^* V_\infty$. With the QR factorization $V_\infty = \mathrm{QR}$, we also have $F = R^* R$, which gives rise to the Cholesky decomposition. Since R has positive diagonal entries, the uniqueness follows. □

3.3 LDL* *decomposition*

As indicated in (3.4), we only need R in order to compute the finite-dimensional representation $\tilde{H}$. The observation made in the preceding section suggests that we may achieve this goal by computing the Cholesky decomposition of F, which can easily be formed as seen in Algorithm 2. However, we notice that when any two roots of $\lambda_1, \dots, \lambda_d$ approach each other (see Example 2 in Section 4), then the matrix F becomes ill-conditioned and indefinite numerically. When this happens, the floating-point errors can easily cause the Cholesky decomposition to break down.

To avoid taking square roots in the intermediate steps of computing the Cholesky decomposition and, more particularly, to exploit its ability in handling indefinite matrices, we suggest the so-called LDL* decomposition as an alternative way of computing R. See Algorithm 3 for the demonstration of LDL* approach. The general theory of LDL* decomposition can be found in (Golub & Van Loan, 2013). When used on indefinite matrices, the LDL* factorization is known to be unstable without careful pivoting. Specifically, similar to the classical *LU* decomposition, the theoretical upper bound of the so-called growth factor for elements in L can be attained for a small class of matrices. A possible improvement is to perform the factorization on block submatrices with special Bunch pivoting. It has been shown that the block $\mathrm{LDL}^{\mathrm{T}}$ factorization in inexact arithmetic is guaranteed to preserve the inertia (Higham, 1999; Fang, 2011). An efficient and high precision implementation is discussed in (Ashcraft *et al.*, 1999) and is available in MATLAB. In particular, a pivoting strategy to improve stability has been employed in the algorithm for the LDL* decomposition. Such a pivoting corresponds to a rearrangement of the values $\lambda_1, \dots, \lambda_d$ in our application.

THEOREM 3.4 There exist a permutation σ of the integers $1, \dots, d$, a unit lower triangular matrix L, and a real diagonal matrix D such that

$$P^{\mathrm{T}} F P = \mathrm{LDL}^*, \tag{3.13}$$

where P is the permutation matrix corresponding to σ and the (k, ℓ) entry of the matrix $P^{\mathrm{T}} F P$ is precisely $f_{\sigma(\ell)}(\bar{\lambda}_{\sigma(k)})$.

Once the LDL* decomposition of F is achieved, we may use the information provided by D to decide whether some extremely small diagonal entries should be discarded, thus modifying the prescribed rank d to a smaller value. This seems to be a reasonable way of estimating the rank. Numerical experiments seem to support this approach.

4. Numerical experiments

In this section, we carry out some numerical experiments with the purpose of demonstrating the theory discussed above. We have already pointed out a few unsettled issues concerning the step from a given Hankel matrix to its Vandermonde factorization. Assuming that hurdle is cleared, our work here is to furnish the finite-dimensional representation when a Vandermonde factorization and the rank d are known. A slightly perturbed low rank matrix is easily of full rank. Thus, in reality, we have to deal with the low rank approximation with Hankel structure only. That is, even though we have a substantial understanding about Hankel operators of finite rank in theory, one prevailing problem is the uncertainty of numerical rank, which is a fairly difficult subject itself in practice. We wish to use examples to bring forth some of the concerns.

EXAMPLE 4.1 We fix $d = 6$ and randomly generate the vibration parameters

$$[\mathbf{a}, \boldsymbol{\zeta}, \boldsymbol{\nu}, \boldsymbol{\phi}] = \begin{bmatrix} 3.8566 & 0.7923 & 0.3948 & 5.7665 \\ 0.1038 & 3.0421 & 51.2192 & 4.4898 \\ 3.1682 & 0.6764 & 81.2621 & 3.4089 \\ 3.7440 & 0.3534 & 61.2526 & 0.8933 \\ 2.4925 & 2.7414 & 72.1755 & 2.3458 \\ 1.1240 & 3.8136 & 29.1876 & 4.2357 \end{bmatrix}$$

whose columns represent the magnitudes a_ℓ, the decay rates ζ_ℓ, the frequencies ν_ℓ and the phase angles ϕ_ℓ, $\ell = 1, \ldots, 6$, for six sinusoids, respectively. For simplicity, we set $\Delta t = 1$. We use the resulting infinite-dimensional Hankel matrix H from the combined signal as our test data. The corresponding Vandermonde factorization (2.9) is readily known. The corresponding numbers λ_ℓ arranged in ascending order of their moduli are

$$0.0084 + 0.0204i, \quad 0.0092 + 0.0468i, 0.0291 + 0.0575i, \quad -0.3575 + 0.2779i,$$
$$-0.0386 + 0.5070i, \quad -0.0115 + 0.7022i.$$

We know that the infinite matrix H should be of rank 6, but we want to compute its finite-dimensional representation $\tilde{H}$ and its singular values.

This problem is small enough that Algorithm 1 is applicable. For curiosity, we plot in Fig. 2 the six orthonormal functions $w_\ell(z)$. These plots of complex-valued functions are to be interpreted as follows. Over the domain of the unit disc (displayed underneath the surface in polar coordinates), the real part $\Re(w_\ell(z))$ of $w_\ell(z)$ is represented by the height of the surface, while the colour of the surface, varying the hue according to the HSV colour model, denotes the imaginary part $\Im(w_\ell(z))$.

We can also use the MATLAB command `ldl` to calculate the LDL* decomposition of the 6×6 sample matrix F to obtain the upper triangular matrix R needed in (3.4). Built in the routine `ldl` is a pivoting strategy which returns that the permutation $\sigma = [1, 6, 4, 5, 3, 2]$. Rearranging the numbers λ_ℓ in accordance with the suggested permutation σ, we regenerate the orthonormal functions $w_\ell(z)$ by Algorithm 1 and the results are plotted in Fig. 3. It is interesting to note that the 'shapes' of these

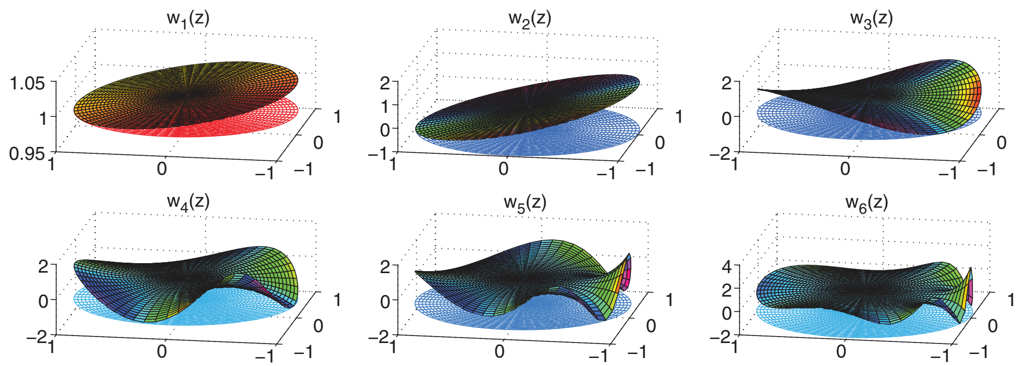

FIG. 2. Sketch of orthonormal functions $w_\ell(z)$ over the unit disc. The height of the surface is the real part, and the colour of the surface is the imaginary part, varying the hue in the HSV colour model.

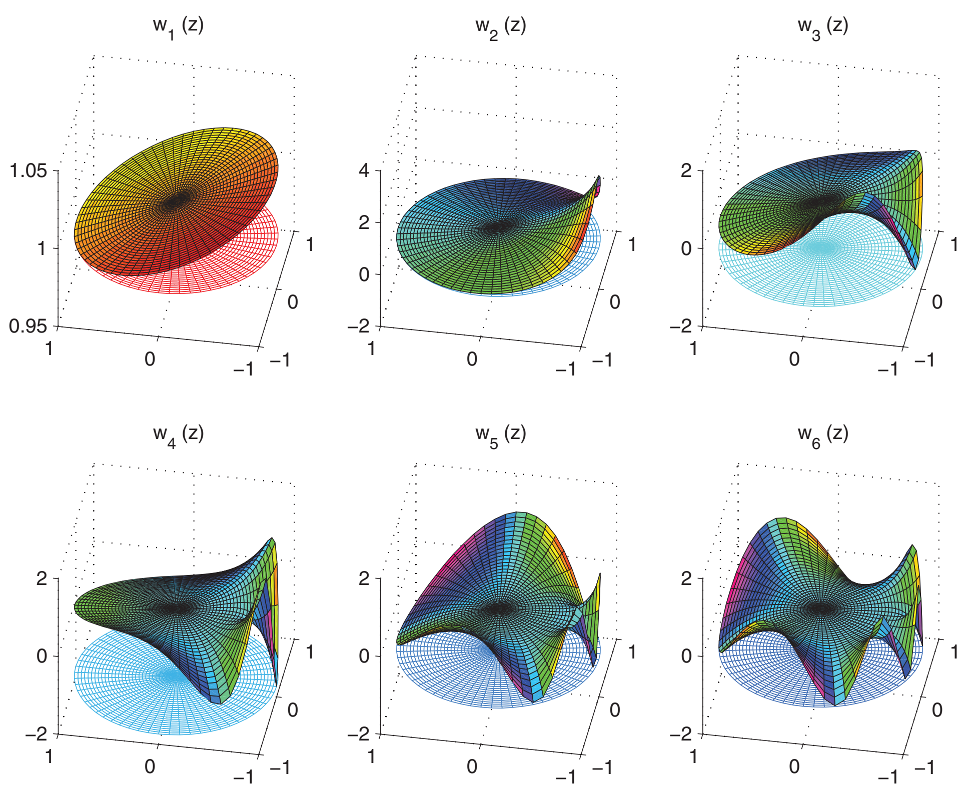

FIG. 3. Sketch of orthonormal functions $w_\ell(z)$ over the unit disc with λ_ℓ rearranged according to LDL* pivoting.

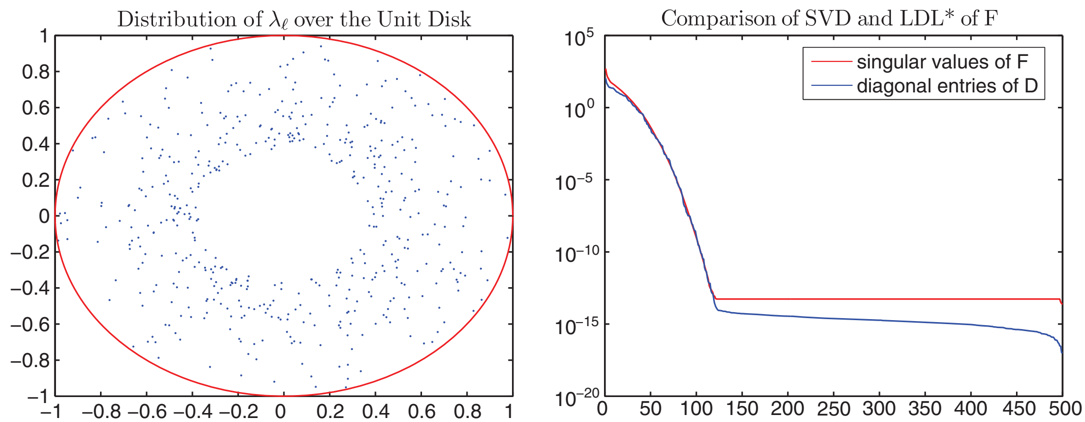

FIG. 4. Comparisons of singular values and diagonal entries of the SVD and the LDL* of the sample matrix F.

orthonormal functions are changed, though their practical meaning is not clear at present. In all cases, the six nonzero singular values of H are found to be

$$4.5999e{+}00,\quad 1.8109e{+}00,\quad 2.1203e{-}01,\quad 1.3692e{-}02,\quad 1.2352e{-}04,\quad 7.4630e{-}10.$$

Even though this test is but a small, arbitrary example, the effect of clustering the numbers λ_ℓ is evident in the smallest singular value. Had we not known that the original signal is composed of six sinusoids beforehand, it would be a close call whether or not to regard the smallest singular value in the above list as nonzero.

EXAMPLE 4.2 The real strength of the LDL* approach is at its efficiency and reliability. First, it is possible to use the Symbolic Math Toolbox in MATLAB to calculate $w_\ell(z)$ via rational arithmetic. But the process is extremely slow. In contrast, the floating-point arithmetic of the MATLAB routine `ldl` applied to the $d \times d$ matrix F is significantly faster. Second, when the rank d is relatively high, we have observed that the value of $r_{\ell\ell}$ obtained by either Algorithms 1 or 2 decreases to nearly zero as ℓ increases, strongly signifying that the modified Gram–Schmidt process is becoming unstable. In contrast, the LDL* calculation handles this stability issue more smoothly.

To demonstrate the advantage of the LDL*, we create a sample matrix F from 500 randomly generated values λ_ℓ over the unit disk. See the left graph in Fig. 4. In theory, F should be of rank 500, and likewise the corresponding Hankel matrix. However, when we plot in the graph of Fig. 4 the singular values of F and the absolute values of the diagonal elements of the matrix D from the LDL* of F, two observations deserve our attention from this experiment. The most noticeable one is that there is a considerable agreement between these two sets of values. The other equally noteworthy observation is that a significant portion of the singular values of F or diagonal elements of D are nearly machine zero. Indeed, numerical experiments indicate that the analogy between singular values of F and the absolute values of diagonal entries of D remains even for low rank d, strongly suggesting that the LDL* handles the nearly singular matrix F almost as well as the singular value decomposition does. These observations are not totally surprising because it is known that the LU algorithms, of which the LDL* is a special case, with suitable pivoting strategies, can reveal ranks properly (Hansen & Yalamov, 2001; Miranian & Gu, 2003; Kawabata *et al.*, 2004; Lee *et al.*, 2009; Foster & Liu, 2012).

With this fact in mind, the detection of nearness to zero by diagonal entries of D suggests that maybe the initial rank d should be reduced. The implication of this rank reduction is significant because it will cause a chain of events to happen. For instance, equation (2.4) will have to be solved for a smaller system, which redefines the coefficients $\gamma_0, \ldots, \gamma_{d-1}$ for the characteristic polynomial (2.10) and, consequently, a new Vandermonde factorization (2.15). Equivalently, a new sample matrix F will be generated of which the LDL* is called again to check the rank condition and the process may be repeated. In terms of the sinusoidal signal (or equivalently its trajectory matrix), rank reduction means that maybe the more complicated initial signal could be approximated with smaller number of components (or a much lower rank Hankel matrix). This rank estimation for Hankel operators, partially related to the subject on SSA, will be reported in a separate paper.

5. Conclusion

Semi-infinite-bounded Hankel matrices of finite rank form an important class of matrices because they occur in applications across multiple fields. This paper makes two contributions. First, on the theoretical side, we revisit existing results and establish connection among the various characterizations of finite rank in terms of rational functions, recursion, sinusoidal signals and Vandermonde factorization. Second, on the computational side, we propose to compute the nonzero singular values of the infinite matrix by means of the LDL* decomposition of a specific finite-dimensional sample matrix.

In all characterizations, though different parameters are used, a semi-infinite-bounded Hankel matrix of rank d depends on $2d$ parameters. The Vandermonde factorization enables us to characterize the action of the original infinite-dimensional H as an operator via a singular-value preserving transformation to a finite-dimensional matrix from which singular values can be calculated.

It is observed that, even though a Hankel matrix is of rank d in exact arithmetic, its numerical rank goes wild as soon as the parameters are replaced by floating-point numbers. The usual modified Gram–Schmidt orthogonalization applied to the generating functions, therefore, becomes increasingly unstable in numerical computation. Under such circumstances, we propose to calculate the upper triangular matrix R needed in the finite-dimensional representation, fast and effectively without any orthogonalization, from the LDL* decomposition of a $d \times d$ sample matrix which is readily available from the Vandermonde factorization. Additionally, the diagonal entries of D might be indicative of rank reduction.

A partial list of future work includes the impact on the singular values when the rank d is wrongly assumed, the effect of multiple or coalescent roots, the sensitivity of the computed roots λ_ℓ subject to perturbations, complexity analysis of the proposed method for singular value computation and the effectiveness of the rank reduction procedure suggested by the LDL* algorithm.

Funding

This research was supported in part by the National Science Foundation under grants DMS-1014166 and DMS-1316779 (M.T.C.). This research was supported in part by the National Science Council of Taiwan under grant NSC 101-2115-M-194-007-MY3 (M.M.L.).

References

Adamjan, V. M., Arov, D. Z. & Kreĭn, M. G. (1968) Infinite Hankel matrices and generalized Carathéodory–Fejér and I. Schur problems. *Funkcional. Anal. i Priložen.*, **2**, 1–17.

AKHIEZER, N. I. (1965) *The Classical Moment Problem and Some Related Questions in Analysis*. New York: Hafner Publishing Co. Translated by N. Kemmer.

ALLEN, M. R. & SMITH, L. A. (1996) Monte Carlo SSA: detecting irregular oscillations in the presence of coloured noise. *J. Climate*, **9**, 3372–3404.

ASHCRAFT, C., GRIMES, R. G. & LEWIS, J. G. (1999) Accurate symmetric indefinite linear equation solvers. *SIAM J. Matrix Anal. Appl.*, **20**, 513–561 (electronic).

AUVERGNE, M. (1988) Singular value analysis applied to phase space reconstruction of pulsating stars. *Astronom. Astrophys.*, **204**, 341–348.

BEZERRA, L. (2012) Vandermonde factorizations of a regular Hankel matrix and their application to the computation of bézier curves. *SIAM J. Matrix Anal. Appl.*, **33**, 411–432.

BOLEY, D. L., LEE, T. J. & LUK, F. T. (1992) The Lanczos algorithm and Hankel matrix factorization. *Linear Algebra Appl.*, **172**, 109–133. Second NIU Conference on Linear Algebra, Numerical Linear Algebra and Applications (DeKalb, IL, 1991).

BOLEY, D. L., LUK, F. T. & VANDEVOORDE, D. (1997) *Vandermonde Factorization of a Hankel Matrix. Scientific Computing (Hong Kong, 1997)*. Singapore: Springer, pp. 27–39.

BOWDEN, C. M. (1968) Boundedness of linear operators in the space ℓ^2. *Int. J. Quantum Chem.*, **2**, 363–371.

BULTHEEL, A. & VAN BAREL, M. (1997) *Linear Algebra, Rational Approximation and Orthogonal Polynomials. Studies in Computational Mathematics*, vol. 6. Amsterdam: North-Holland Publishing Co.

CHU, M. T., FUNDERLIC, R. E. & PLEMMONS, R. J. (2003) Structured low rank approximation. *Linear Algebra Appl.*, **366**, 157–172. Special issue on Structured Matrices: Analysis, Algorithms and Applications (Cortona, 2000).

DANILOV, D. L. (1997) Principal components in time series forecast. *J. Comput. Graph. Statist.*, **6**, 112–121.

DUTOIT, T. (2004) Unusual teaching short-cuts to the Levinson and lattice algorithms. *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)*, vol. 5. Quebec: IEEE, pp. V–1029–32.

FANG, H.-R. (2011) Stability analysis of block ldl^{T} factorization for symmetric indefinite matrices. *IMA J. Numer. Anal.*, **31**, 528–555.

FELDMANN, S. & HEINIG, G. (1996) Vandermonde factorization and canonical representations of block Hankel matrices. *Proceedings of the Fourth Conference of The International Linear Algebra Society*, vol. 241/243. Rotterdam: Elsevier, pp. 247–278.

FOSTER, L. V. & LIU, X. (2012) Comparison of rank revealing algorithms applied to matrices with well defined numerical rank. *Technical Report*. San Jose: San Jose State University. Notes and MATLAB codes available at http://www.math.sjsu.edu/ foster/rankrevealingcode.html.

FRANCIS, B. A. (1987) *A Course in H_∞ Control Theory*. Berlin: Springer.

GANTMACHER, F. R. (1959) *The Theory of Matrices. Vols. 1, 2*. New York: Chelsea Publishing Co. Translated by K. A. Hirsch.

GHIL, M. & TARICCO, C. (1997) Advanced spectral analysis methods. *Past and Present Variability of the Solar-Terrestrial System: Measurement, Data Analysis and Theoretical Models* (G. C. Castagnoli & A. Provenzale eds). Bologna, Italy: Soc. Ital. di Fis., pp. 137–159.

GILLARD, J. & ZHIGLJAVSKY, A. (2011) Analysis of structured low rank approximation as an optimization problem. *Informatica (Vilnius)*, **22**, 489–505.

GLOVER, K. (1984) All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. *Internat. J. Control*, **39**, 1115–1193.

GOLUB, G. H. & VAN LOAN, C. F. (2013) *Matrix Computations*, 4th edn. Baltimore: Johns Hopkins University Press.

GOLUB, G. H. & WELSCH, J. H. (1969) Calculation of Gauss quadrature rules. *Math. Comp.*, **23**, 221–230.

GOLYANDINA, N., NEKRUTKIN, V. & ZHIGLJAVSKY, A. (2001) *Analysis of Time Series Structure: SSA and Related Techniques*. Boca Raton: Chapman & Hall/CRC.

GONNET, P., GÜTTEL, S. & TREFETHEN, L. (2013) Robust Padé approximation via SVD. *SIAM Review*, **55**, 101–117.

Gonzalez, R. C. & Woods, R. E. (2008) *Digital Image Processing*, 3rd edn. New Jersey: Prentice Hall.

Gragg, W. B. & Reichel, L. (1989) On singular values of Hankel operators of finite rank. *Linear Algebra Appl.*, **121**, 53–70.

Hansen, P. C. & Yalamov, P. Y. (2001) Computing symmetric rank-revealing decompositions via triangular factorization. *SIAM J. Matrix Anal. Appl.*, **23**, 443–458 (electronic).

Heinig, G. (2001) Fast and superfast algorithms for Hankel-like matrices related to orthogonal polynomials. *Numerical Analysis and Its Applications (Rousse, 2000)*. Lecture Notes in Computer Science, vol. 1988. Berlin: Springer, pp. 385–392.

Heinig, G. & Rost, K. (1984) *Algebraic Methods for Toeplitz-Like Matrices and Operators*. Mathematical Research, vol. 19. Berlin: Akademie-Verlag.

Higham, N. J. (1999) Stability of block LDL^{T} factorization of a symmetric tridiagonal matrix. *Linear Algebra Appl.*, **287**, 181–189. Special issue celebrating the 60th birthday of Ludwig Elsner.

Ibryaeva, O. L. & Adukov, V. M. (2013) An algorithm for computing a Padé approximant with minimal degree denominator. *J. Comput. Appl. Math.*, **237**, 529–541.

Iohvidov, I. S. (1982) *Hankel and Toeplitz Matrices and Forms*. Boston, MA: Birkhäuser. Translated by G. Philip A. Thijsse.

Kalman, D. (1984) The generalized Vandermonde matrix. *Math. Mag.*, **57**, 15–21.

Kawabata, Y., Matsubara, M. & Sugimoto, S. (2004) System order determination by using $\mathrm{LDL}^{\mathrm{T}}$ decomposition and information criteria. *Proceedings of the 35th ISCIE International Symposium on Stochastic Systems Theory and its Applications*. Ube, Japan: Inst. Syst. Control Inform. Engrs. (ISCIE), Kyoto, pp. 158–163.

Kibangou, A. Y. & Favier, G. (2007) Toeplitz-Vandermonde matrix factorization with application to parameter estimation of Wiener-Hammerstein systems. *Signal Process. Lett., IEEE*, **14**, 141–144.

Kung, S. Y. (1978) A new identification and model reduction algorithm via singular value decomposition. *Proceedings of the 12th Asilomar Conference on Circuits, Systems and Computers*. Pacific Grove: IEEE, pp. 705–714.

Kung, S. Y. & Lin, D. W. (1981) Recent progress in linear system model-reduction via Hankel matrix approximation. *Circuit Theory and Design (The Hague, 1981)* (R. Boite & P. Dewilde eds). Amsterdam: North-Holland, pp. 222–233.

Lee, T.-L., Li, T.-Y. & Zeng, Z. (2009) A rank-revealing method with updating, downdating, and applications. II. *SIAM J. Matrix Anal. Appl.*, **31**, 503–525.

Lemmerling, P. & Van Huffel, S. (2001) Analysis of the structured total least squares problem for Hankel/Toeplitz matrices. *Numer. Algorithms*, **27**, 89–114.

Luk, F. T. & Qiao, S. (2003) A fast singular value algorithm for Hankel matrices. *Fast Algorithms for Structured Matrices: Theory and Applications*, vol. 323. Providence: Amer. Math. Soc., pp. 169–177.

Markovsky, I. (2012) *Low Rank Approximation: Algorithms, Implementation, Applications*. London: Springer.

Markovsky, I., Willems, J. C., Van Huffel, S., De Moor, B. & Pintelon, R. (2005) Application of structured total least squares for system identification and model reduction. *IEEE Trans. Automat. Control*, **50**, 1490–1500.

Mineva, A. & Popivanov, D. (1996) Method for single-trial readiness potential identification, based on singular spectrum analysis. *J. Neurosci. Methods*, **68**, 91–99.

Miranian, L. & Gu, M. (2003) Strong rank revealing *LU* factorizations. *Linear Algebra Appl.*, **367**, 1–16.

Nehari, Z. (1957) On bounded bilinear forms. *Ann. Math.* (2), **65**, 153–162.

Pan, V. Y. (2001) *Structured Matrices and Polynomials: Unified Superfast Algorithms*. Boston: Birkhäuser.

Park, H., Zhang, L. & Rosen, J. B. (1999) Low rank approximation of a Hankel matrix by structured total least norm. *BIT*, **39**, 757–779.

Partington, J. R. (1988) *An Introduction to Hankel Operators*. Cambridge: Cambridge University Press.

Peller, V. V. (1998) An excursion into the theory of Hankel operators. *Holomorphic Spaces (Berkeley, 1995)* (S. Axler, J. E. McCarthy, & D. Sarason eds). Cambridge: Cambridge Univ. Press, pp. 65–120.

Peller, V. V. (2003) *Hankel Operators and their Applications*. New York: Springer.

Power, S. C. (1982) *Hankel Operators on Hilbert Space*. Boston: Pitman (Advanced Publishing Program).

STEELE, J. M. (2012) The Yule–Walker equations and algorithms for their solution. Notes and references available at http://www-stat.wharton.upenn.edu/ steele/Courses/956/ResourceDetails/Yul eWalkerAndMore.htm,.

TYRTYSHNIKOV, E. (2010) Hankel minors and Pade approximations. *Numerical Methods for Structured Matrices and Applications*. Operator Theory: Advances and Applications (D. A. Bini, V. Mehrmann, V. Olshevsky, E. Tyrtsyhnikov & M. Van Barel eds), vol. 199. Basel: Birkhäuser Verlag, pp. 431–439.

VAUTARD, R., YIOU, P. & GHIL, M. (1992) Singular-spectrum analysis: a toolkit for short, noisy chaotic signals. *Phys. D*, **58**, 95–126.

WIDOM, H. (1966) Hankel matrices. *Trans. Amer. Math. Soc.*, **121**, 1–35.

YIOU, P., SORNETTE, D. & GHIL, M. (2000) Data-adaptive wavelets and multi-scale singular-spectrum analysis. *Phys. D*, **142**, 254–290.

YOUNG, N. J. (1983) The singular-value decomposition of an infinite Hankel matrix. *Linear Algebra Appl.*, **50**, 639–656.