The Singular-Value Decomposition of an Infinite Hankel Matrix

N. J. Young
Department of Mathematics
University Gardens
University of Glasgow
Glasgow, G12 8QW, Scotland

ABSTRACT

Let H be an infinite Hankel matrix of known finite rank r. A new algorithm for the numerical calculation of the singular values and vectors of H is presented. The method proceeds by reduction to the singular value problem for an $r \times r$ matrix; this is achieved without solving for the poles of the symbol of H. The resulting algorithm is of order r^3.

1. INTRODUCTION

Hankel operators are becoming popular in various aspects of engineering design. Their connection with moment problems has been known for a hundred and fifty years, but it is the recent results of V. Adamyan, D. Arov, and M. Krein [1] relating them to norm-approximation problems which have provided the impetus for new applications. These include digital-filter design [7], model reduction [10], and broadband matching [8]. In each of these cases the calculation of a transfer function or impedance with some desired optimality property is reduced to the determination of certain singular values and vectors of an infinite Hankel matrix: that is, a matrix $H = [c_{i+j}]_{i,j=0}^{\infty}$ with $c_i \in \mathbb{C}$.

The purpose of this paper is to describe a new algorithm for this singular-value problem. This algorithm has been implemented on a computer and has performed most satisfactorily in tests, being fast and stable.

The most obvious way to tackle the singular-value problem for an infinite matrix is simply to truncate and thereby reduce to a finite-dimensional problem, but this is not necessarily the best approach. Consider the Hankel

52 Vanderbilt Ave., New York, NY 10017 0024-3795/83/020639-18$3.00

matrix

$$H = \begin{bmatrix} 1 & a & a^2 & \cdots \\ a & a^2 & a^3 & \cdots \\ a^2 & a^3 & a^4 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

where $0 < a < 1$. H has rank one, and its unique nonzero singular value is easily seen to be $(1 - a^2)^{-1}$. If we truncate after k rows and columns, we obtain a rank-one matrix whose nonzero singular value is $(1 - a^{2k})/(1 - a^2)$, so that the truncation error is $a^{2k}/(1 - a^2)$. If, say, $a = 0.9$, we must take k to be around 37 for this error to be less than 10^{-3}. To get three-figure accuracy we must therefore operate with 37×37 matrices, despite the fact that H can be represented by a 1×1 matrix. In practice things would be even worse: one would not know in advance the right value of k, and would therefore have to solve the singular-value problem for a succession of increasingly large truncations in order to observe convergence. In the alternative method described here, if H is of rank r, then one operates throughout with $r \times r$ matrices which in principle represent the operator exactly and introduce no truncation error. The only iteration involved is in solving the singular-value problem for an $r \times r$ matrix. The method depends, of course, on H having finite rank and this rank (or a reasonable upper bound therefore) being known. These requirements will often be met in practice, since the data from which H is generated will have been obtained by a finite number of measurements. Even if H is not of known finite rank, one would attempt to apply the algorithm; we consider this possibility briefly in Section 7 below.

Alternatives to the present algorithm have been put forward by Sun-Yan Kung [10] and Ph. Delsarte, Y. Genin, and Y. Kamp [4]. These are discussed in Section 6.

A caution as to the application of this method is needed. Adamyan, Arov, and Krein enable us to reduce some important problems to the singular-value problem for Hankel matrices, but I believe that in many cases it is better not to do so. Take, for example, the Nevalinna-Pick problem [1,4]: the AAK reduction yields a Hankel matrix whose entries are the Taylor coefficients of a certain rational function. If the present method were applied to this Hankel matrix, the first step would be to recover the rational function, with necessarily some loss of accuracy. Folklore in nuclear physics supports this view: approximation methods which rely on moments are held to be numerically unsatisfactory (see the introduction to [13]). Programs which apply directly to various forms of the Nevanlinna-Pick problem (and hence to certain design problems [4]) are being developed by Dr. A. C. Allison and the author and will be described elsewhere. These programs are based on the "dual extremal" approach of D. Sarason [12] rather than that of Adamyan, Arov, and Krein. Theorem 3 below essentially describes the connection between these two approaches.

Let us establish some notation. We write l^2 for the Hilbert space of semiinfinite square summable sequences,

$$l^2 = \left\{ (x_j)_0^\infty : x_j \in \mathbb{C},\ \sum_0^\infty |x_j|^2 < \infty \right\},$$

with the inner product

$$\left[(x_j), (y_j) \right] = \sum_0^\infty x_j \bar{y}_j$$

(bars always denote complex conjugation). An infinite matrix $T = [t_{ij}]_{i,j=0}^\infty$ will be said to act on l^2 if it defines a bounded linear operator on l^2 in the obvious way. If an infinite matrix T acts on l^2 and has finite rank r, then there exist positive numbers $s_0 \geqslant s_1 \geqslant \cdots \geqslant s_{r-1}$ and orthonormal sequences $e_0, \dots, e_{r-1}$ and $f_0, \dots, f_{r-1}$ in l^2 such that

$$T = \sum_{j=0}^{r-1} s_j(\cdot, f_j) e_j \tag{1}$$

in the sense that, for any $x \in l^2$,

$$Tx = \sum_{j=0}^{r-1} s_j(x, f_j) e_j.$$

The s_j are called the *singular values* of T and are unique, being the nonzero eigenvalues of $(T^*T)^{1/2}$ (T^* denotes the adjoint operator of T); 0 is also admitted as a singular value. e_j and f_j are called *singular vectors*, and the ordered pair $\langle e_j, f_j \rangle$ is called a *Schmidt pair for* T corresponding to the singular value s_j. The expression (1) is called the *singular-value decomposition of* T. Properly speaking one should use the indefinite article, since it is not unique in general; however, if s_j is a simple eigenvalue of $(T^*T)^{1/2}$, then $\langle e_j, f_j \rangle$ is determined up to multiplication by a scalar of unit modulus.

2. REDUCTION TO FINITE DIMENSIONS

Consider an infinite Hankel matrix H acting on l^2,

$$H = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots \\ c_1 & c_2 & c_3 & \cdots \\ c_2 & c_3 & c_4 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

of finite rank r. The range (column space) R of H is then, by the definition of rank, of dimension r. It follows that the singular-value problem for H can be reduced to the corresponding problem for an $r \times r$ matrix. Indeed, if T is any operator on a Hilbert space, then we can regard T as the orthogonal direct sum of the zero operator from $\operatorname{Ker} T$ into $(\operatorname{Range} T)^\perp$ and the restriction of T acting from $(\operatorname{Ker} T)^\perp$ into $\operatorname{Range} T$. In the case $T = H$ we have

$$(\operatorname{Ker} H)^\perp = \operatorname{Range} H^* = \operatorname{Range} \bar{H} = \tilde{R},$$

where $\tilde{x}$ denotes the conjugate of x [if $x = (x_0, x_1, \dots)$ then $\tilde{x} = (\bar{x}_0, \bar{x}_1, \dots)$] and $\tilde{R} = \{\tilde{x} : x \in R\}$. Thus H is the orthogonal direct sum of a zero operator and an operator between two r-dimensional spaces, $\tilde{R}$ and R.

Let us now convert to an $r \times r$ matrix problem. The calculations can be carried out conveniently in terms of generating functions. As is customary, we denote by H^2 the space of generating functions of l^2 sequences:

$$H^2 = \left\{ \sum_0^\infty x_j z^j : \sum_0^\infty |x_j|^2 < \infty \right\}.$$

H^2 is a Hilbert space of functions analytic in the open unit disc (see [9] or [5] for a fuller discussion). Let us write

$$\psi(z) = \sum_0^\infty c_n z^n,$$

and note that since H is the matrix of an operator on Hilbert space, its columns are square summable, and so $\psi \in H^2$. Furthermore, the fact that H has finite rank implies that ψ is a rational function (this is a well-known theorem of Kronecker, and will in any case emerge from calculations below).

It is easily shown that a rational H^2 function can have no poles on the unit circle, so all the poles of ψ lie outside the closed unit disc. Thus $\psi \in H^\infty$, the subspace of H^2 consisting of those functions essentially bounded on the unit circle. We denote by T_ψ the operator on H^2 whose matrix with respect to the standard orthonormal basis $1, z, z^2, \dots$ is H.

We also need the space L^2 of square integrable functions on the unit circle: recall that, by the Fischer-Riesz theorem,

$$L^2 = \left\{ \sum_{-\infty}^{\infty} x_j z^j : \sum_{-\infty}^{\infty} |x_j|^2 < \infty \right\}.$$

Thus H^2 can be identified with the subspace of L^2 comprising those functions whose negative Fourier coefficients vanish, and we may therefore introduce the orthogonal projection operator $P: L^2 \to H^2$. Then for $j = 0, 1, \dots,$

$$\begin{aligned} T_\psi z^j &= c_j + c_{j+1} z + c_{j+2} z^2 + \cdots \\ &= P\left(c_0 \bar{z}^j + c_1 \bar{z}^{j-1} + \cdots + c_j + c_{j+1} z + \cdots \right) \\ &= P\left(\bar{z}^j \psi(z) \right) \\ &= P M_\psi J z^j, \end{aligned}$$

where J is the "reversal operator" on L^2 [that is, $Jf(z) = f(\bar{z})$] and M_ψ is the operator on L^2 of multiplication by $\psi(z)$. Hence

$$T_\psi = P M_\psi J | H^2. \tag{2}$$

The next step is to describe Range T_ψ and $(\operatorname{Ker} T_\psi)^\perp$. To this end let us see how to obtain a rational expression for ψ. Observe that the generating function of the jth column of H is

$$\frac{1}{z^j}\left[\psi(z) - c_0 - c_1 z - \cdots - c_{r-1} z^{j-1} \right].$$

As H has rank r, the first $r+1$ columns are linearly dependent and so there exist $a_0, \dots, a_r \in \mathbb{C}$, not all zero, such that

$$a_0 \psi(z) + \frac{a_1}{z}\left[\psi(z) - c_0 \right] + \cdots + \frac{a_r}{z^r}\left[\psi(z) - c_0 - \cdots - c_{r-1} z^{r-1} \right] = 0, \tag{3}$$

which can be solved to give

$$\psi(z)=\frac{g(z)}{q(z)}, \tag{4}$$

where g is a polynomial of degree not exceeding $r-1$ and

$$q(z)=a_0z^r+a_1z^{r-1}+\cdots+a_r. \tag{5}$$

THEOREM 1. *If ψ is given by* (4), *and T_ψ has rank $r\geqslant 1$, then* Range T_ψ *is the space of all functions of the form f/q where f is a polynomial of degree less than r.*

We shall make use of the *backward shift operator* S on H^2: this is defined by

$$S(x_0+x_1z+x_2z^2+\cdots)=x_1+x_2z+x_3z^2+\cdots, \tag{6}$$

or equivalently

$$Sf(z)=\begin{cases}\frac{1}{z}\{f(z)-f(0)\} & \text{if} \quad z\neq 0,\\ f'(0) & \text{if} \quad z=0,\end{cases}$$

f here being regarded as a function on the open unit disc.

Proof of Theorem 1. Let E be the space of rational functions f/q with q given by (5), and f a polynomial of degree $<r$. Observe that

$$S(f/q)(z)=\frac{f(z)q(0)-f(0)q(z)}{zq(z)q(0)},$$

so that E is invariant under S (i.e. $SE\subseteq E$). Now $\psi\in E$, and the generating functions of the columns of H are $\psi, S\psi, S^2\psi,\ldots,$ which is to say that $T_\psi z^j=S^j\psi$. Hence Range $T_\psi\subseteq E$. Since both spaces have dimension r, we must have Range $T_\psi=E$, as required. ∎

The space of rational functions described in Theorem 1 plays an important role: let us denote it by K_q. Further, we denote the natural conjugation operator on H^2 by C:

$$C(x_0 + x_1 z + \cdots) = \bar{x}_0 + \bar{x}_1 z + \cdots,$$

or equivalently

$$Cf(z) = f(\bar{z})^{-}.$$

C is a conjugate-linear operator.

THEOREM 2. $(\operatorname{Ker} T_\psi)^\perp = K_{Cq}$.

Proof.

$$\begin{aligned}(\operatorname{Ker} T_\psi)^\perp &= \text{closure of Range } T_\psi^* \\ &= \text{Range } T_\psi^*,\end{aligned}$$

since the latter space is finite dimensional. Now $T_\psi^* = T_{C\psi}$: this is immediate from the fact that $H^* = \bar{H}$. Since $C\psi = Cg/Cq$, Theorem 2 now follows from Theorem 1. ∎

Let V_ψ denote the restriction of T_ψ mapping K_{Cq} into K_q; then T_ψ is the orthogonal direct sum of a zero operator and V_ψ. We could now attempt to reduce to a matrix problem by choosing bases in K_{Cq} and K_q, but the calculations seem to be easier if we first relate V_ψ to an operator acting on K_q.

THEOREM 3. *The diagram*

$$\begin{array}{ccc} & K_{Cq} & \\ {}^{U}\swarrow & & \searrow^{V} \\ K_q & \xrightarrow[h(S)]{} & K_q \end{array}$$

commutes, where U is defined by

$$\left(U\frac{f}{Cq}\right)(z) = \frac{z^{r-1} f(\bar{z})}{q(z)}$$

and

$$h(z)=\frac{z^{r-1}g(\bar{z})}{Cq(z)}.$$

Proof. Note first that $h(S)$ is well defined, since h is analytic in a neighborhood of the closed unit disc while S has norm 1, and hence has spectrum contained in the closed unit disc. Another way of writing the definition (6) of S is

$$Sf(z)=P(\bar{z}f(z)),$$

from which it follows that

$$h(S)f(z)=P(h(\bar{z})f(z))$$

(for example by using the power-series expansion of h). Thus, for an arbitrary element f/Cq of K_{Cq} we have

$$\begin{aligned}\left(h(S)U\frac{f}{Cq}\right)(z)&=h(S)\frac{z^{r-1}f(\bar{z})}{q(z)}\\&=P\left(\frac{\bar{z}^{r-1}g(z)}{Cq(\bar{z})}\frac{z^{r-1}f(\bar{z})}{q(z)}\right)\\&=P\left(\frac{g(z)}{q(z)}\frac{f(\bar{z})}{Cq(\bar{z})}\right)=PM_{\psi}J\frac{f}{Cq}(z)\\&=T_{\psi}\frac{f}{Cq}(z)=V_{\psi}\frac{f}{Cq}(z).\end{aligned}$$ ■

Note that U is unitary mapping of K_{Cq} onto K_q: in fact $|Uf(z)|=|f(\bar{z})|$ for all $f\in K_{Cq}$ and $|z|=1$. Thus the singular value problem for V_{ψ} is equivalent to that for $h(S)$. The advantage of this is that the matrix of $h(S)$ is easier to calculate, for since K_q is invariant under S,

$$h(S)|K_q=h(S|K_q).$$

The remaining steps are to choose an orthonormal basis for K_q relative to which $S|K_q$ has an easily calculated matrix M, to form $h(M)$, and to invoke a standard routine to solve the singular-value problem for $h(M)$.

Once we have found the singular values and vectors of $h(S)|K_q$, it is of course easy to obtain those of V_ψ. Suppose

$$h(S)|K_q = \sum_{i=0}^{r-1} s_i(\cdot, y_i)x_i.$$

Then, for any $f \in K_{Cq}$ we have

$$\begin{aligned} V_\psi f &= h(S)Uf \\ &= \sum_{i=0}^{r-1} s_i(Uf, y_i)x_i, \end{aligned}$$

and so the singular-value decomposition of V_ψ is

$$V_\psi = \sum_{i=0}^{r-1} s_i(\cdot, U^* y_i)x_i. \tag{7}$$

It is easy to check that, for any function $f/q \in K_q$,

$$\left(U^* \frac{f}{q}\right)(z) = \left(U^{-1}\frac{f}{q}\right)(z) = \frac{z^{r-1}f(\bar{z})}{Cq(z)}. \tag{8}$$

Incidentally, this shows that $h = U^*\psi$.

3. A USEFUL BASIS

The preceding section shows that we shall be able to find the singular values and vectors of the Hankel matrix H provided we can choose an orthonormal basis of the space K_q of rational functions with respect to which the restricted shift operator $S|K_q$ has an easily calculated matrix. In the case that the zeros of the polynomial q are known it is not hard to write down such a basis, and so we could proceed by first solving the equation $q(z) = 0$. However, it is well known that this is liable to lead to numerical instability, and so it is fortunate that there is an alternative method which avoids the solution of a polynomial equation. There is a very natural basis of K_q, expressible in terms of the coefficients of q rather than its zeros, which has two great advantages: firstly, the matrix with respect to this basis of $S|K_q$ is a companion matrix. This is vital to the efficiency of the algorithm, as it enables the computation of the matrix of $h(S|K_q)$ with an operation count of $O(r^3)$, rather than the $O(r^4)$ one might expect. The second advantage is that the basis, though not orthonormal, can be readily orthogonalized, since there is a neat and simple formula for its Gram matrix. It is this formula which is the technical innovation underlying the present algorithm and the related ones for the Nevanlinna-Pick problem which have been developed by Dr. A. C. Allison and the author.

The basis in question consists of the functions $f_0,\dots,f_{r-1}$, where f_j is the unique member of K_q having a Taylor series of the form

$$f_j(z)=z^j+O(z^r). \tag{9}$$

There is in fact a unique f_j of the form b_j/q, with b_j of degree less than r, satisfying (9), for (9) is equivalent to

$$b_j(z)=z^j(a_r+a_{r-1}z+\cdots+a_0z^r)+z^rq(z)k(z)$$

for some $k\in H^2$, and this is equivalent to

$$b_j(z)=a_rz^j+a_{r-1}z^{j+1}+\cdots+a_{j+1}z^{r-1}.$$

Now

$$Sf_j(z)=\frac{b_j(z)-\dfrac{b_j(0)}{a_r}q(z)}{zq(z)},$$

and so, for $j=1,\dots,r-1$, we have $b_j(0)=0$ and

$$\begin{aligned}Sf_j(z)&=\frac{a_rz^{j-1}+\cdots+a_{j+1}z^{r-2}}{q(z)}\\&=f_{j-1}(z)-a_jz^{r-1}/q(z)\\&=f_{j-1}(z)-\frac{a_j}{a_r}f_{r-1}(z),\end{aligned} \tag{10}$$

while

$$
\begin{aligned}
Sf_0(z) &= \frac{b_0(z) - q(z)}{zq(z)} \\
&= -a_0 z^{r-1}/q(z) \\
&= -\frac{a_0}{a_r} f_{r-1}(z).
\end{aligned} \tag{11}
$$

Putting (10) and (11) together, we infer that the matrix of $S|K_q$ with respect to $f_0, \ldots, f_{r-1}$ is

$$
F = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
-a_0/a_r & -a_1/a_r & -a_2/a_r & \cdots & -a_{r-1}/a_r
\end{bmatrix}
$$

The latter matrix is called the *companion* or *Frobenius matrix* of the polynomial

$$
\tilde{q}(z) = a_0 + a_1 z + \cdots + a_r z^r. \tag{12}
$$

The matrix of $h(S)|K_q$ with respect to $f_0, \ldots, f_{r-1}$ is now seen to be $h(F)$.

To obtain the matrix of this operator with respect to an orthonormal basis we use standard facts about change of basis which we now summarize.

LEMMA. *Let A, Q be linear transformations on K_q having matrices A_f, Q_f respectively with respect to $f_0, \ldots, f_{r-1}$.*

(*i*) *If Q is invertible, the matrix of A with respect to $Qf_0, \ldots, Qf_{r-1}$ is $Q_f^{-1} A Q_f$;*

(ii) *$Qf_0, \ldots, Qf_{r-1}$ is an orthonormal basis if and only if $Q_f Q_f^* = G^{-1}$ when G is the Gram matrix of $f_0, \ldots, f_{r-1}$,* i.e. $G = [g_{ij}]$, *where* $g_{ij} = (f_j, f_i)$.

The formula for the Gram matrix G of $f_0, \ldots, f_{n-1}$ mentioned above is as follows.

THEOREM 4. *The Gram matrix G of the basis $f_0,\dots,f_{n-1}$ of K_q is given by*

$$G=(I-B^*B)^{-1}, \tag{13}$$

where

$$B=\begin{bmatrix}\beta_1 & \beta_2 & \beta_3 & \cdots & \beta_r\\ 0 & \beta_1 & \beta_2 & \cdots & \beta_{r-1}\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & \beta_1\end{bmatrix}$$

and

$$\bar{a}_r\beta_1=a_0,$$

$$\bar{a}_r\beta_i=a_{i-1}-\bar{a}_{r-i+1}\beta_1-\bar{a}_{r-i+2}\beta_2-\cdots-\bar{a}_{r-1}\beta_{i-1}$$

for $2\leqslant i\leqslant r$.

This formula is proved in [14, Remark 4]. It depends on the characterization of G as the unique solution of the matrix equation

$$G-F^*GF=\operatorname{diag}\{1,0,\dots,0\}.$$

The formula (13) enables us to compute G^{-1} with only r^2 multiplications and divisions. To apply Lemma 4 we simply perform a Choleski decomposition of G^{-1}—that is, find a triangular matrix Q_f such that $Q_fQ_f^*=G^{-1}$—and form $Q_f^{-1}h(F)Q_f$, which will be the matrix of $h(S)|K_q$ with respect to the orthonormal basis $Qf_0,\dots,Qf_{r-1}$.

4. THE ALGORITHM

The previous two sections contain the necessary ingredients for a recipe for the solution of the singular-value problem for the Hankel matrix H. The algorithm proceeds as follows.

1. Read in the rank r and $2r$ entries $c_0,\dots,c_{2r-1}$ of H.
2. Find a linear relation between the first $r+1$ columns of H: that is, find $a_0,\dots,a_r$ of Equation (3).

3. Form the polynomials g and q such that

$$\frac{g(z)}{q(z)} = \psi(z) = \sum_{0}^{\infty} c_n z^n.$$

[Equation (4)].
4. Form the rational function $h(z) = z^{r-1}g(\bar{z})/Cq(z)$ (see Theorem 3).
5. Form the matrix $h(F)$, where F is the companion matrix of the polynomial $\tilde{q}$ [equation (12)].
6. Evaluate the inverse Gram matrix G^{-1} according to the formula (13).
7. Perform a Choleski decomposition to obtain a triangular matrix Q_f such that $Q_f Q_f^* = G^{-1}$.
8. Form the matrix $A = Q_f^{-1} h(F) Q_f$, which is the matrix of $h(S)|K_q$ with respect to the orthonormal basis $Qf_0, \dots, Qf_{r-1}$ (see Lemma).
9. Invoke a library routine to obtain the singular values $s_0, \dots, s_{r-1}$ of A and the corresponding Schmidt pairs $\langle u_0, v_0 \rangle, \dots, \langle u_{r-1}, v_{r-1} \rangle$.
10. Print $s_0, \dots, s_{r-1}$, which are the desired singular values of H, all other singular values being zero.
11. For $i = 1, \dots, r$ obtain the Schmidt pair $\langle x_i, y_i \rangle$ of $h(S)|K_q$ as the rational functions in K_q whose components with respect to $Qf_0, \dots, Qf_{r-1}$ are given by u_i, v_i respectively. Thus, if $Q_f u_i = (\xi_0, \dots, \xi_{r-1})$ and $Q_f v_i = (\eta_0, \dots, \eta_{r-1})$, we have

$$x_i = \xi_0 f_0 + \cdots + \xi_{r-1} f_{r-1}, \qquad y_i = \eta_0 f_0 + \cdots + \eta_{r-1} f_{r-1}.$$

12. Calculate and print the Schmidt pairs $\langle x_i, U^* y_i \rangle$ of T_ψ using Equation (8).

The rational functions x_i and $U^* y_i$ are of course the generating functions of the l^2 sequences which constitute the Schmidt pair of H corresponding to the singular value s_i.

5. COMMENTS ON THE ALGORITHM

Some of the above steps require further discussion. Steps 2 and 3 require us to find the rational function $\psi(z) = \sum_0^\infty c_n z^n$ from a knowledge of the integer r (which, by Kronecker's theorem, is the degree of q) and the $2r$ entries $c_0, \dots, c_{2r-1}$. This is an instance of a long-established procedure — the derivation of the $(r-1, r)$ *Padé approximant* of ψ (which here actually coincides with ψ). There are standard methods which achieve this task in

$O(r^2)$ operations, while a recent development is to use the fast Fourier transform to reduce the operation count to $O(r\log^2 r)$ (see [2]). However, these fast methods aie known to carry some risk of instability, and since the two main stages of the present algorithm (evaluating a function of a companion matrix and solving a singular-value problem) are both of order r^3, there is no point in economizing on operations in the Padé approximation. Accordingly my own implementation effects this step using the QR decomposition, which is of order r^3 but is more reliable.

Finding the rational function ψ is equivalent to solving a homogeneous system of linear equations, and this might turn out to be ill conditioned. The necessity for this step is the main drawback of the present algorithm in comparison with the more obvious method of truncation. Whether this possible source of instability is sufficiently compensated by the reduction in size of the singular-value problem remains to be decided by experience. Practical tests to date seem quite favorable. In the event that the rational function ψ is known in advance (as is assumed in [10]), the Padé step is unnecessary, making the present method considerably more competitive.

There are two further steps of numerical significance. One is step 5: the evaluation of $h(F)$ where h is a rational function g/Cq and F is the companion matrix of $\tilde{q}$. This depends on the Euclidean algorithm. Provided that H is bounded (without which assumption the whole problem is meaningless), the zeros of $\tilde{q}$ lie in the open unit disc while those of Cq lie outside it, so that Cq and $\tilde{q}$ are relatively prime polynomials. We can therefore use the Euclidean algorithm to find polynomials u and v such that

$$uCq + v\tilde{q} = 1.$$

Then u is a reciprocal of Cq (mod $\tilde{q}$) and so $Cq(F)^{-1} = u(F)$. Hence to evaluate $h(F)$ it suffices to find the remainder f on dividing $\hat{g}u$ by $\tilde{q}$, which is a polynomial of degree less than r, and to use the fact that $h(F) = f(F)$. Now it is easy to see from the form of the companion matrix F that when one takes powers $F^2, F^3, \dots,$ one introduces only one new row at each step, the bottom $r-1$ rows being shifted up. Using this fact we can compute $f(F)$ in about $\frac{1}{2}r^3$ operations.

Since we have such a simple formula for G^{-1}, step 6 is safe and easy, but 7 and 8 could present dangers. In fact it follows from [15, Equation (29)] and [14, Remark 2] that G^{-1} has determinant

$$\prod_{i,j=1}^{r} (1 - p_j\bar{p}_i),$$

where $p_1,\dots,p_r$ are the zeros of $\tilde{q}$. This suggests very strongly that steps 7 and 8 will tend to produce instability precisely when some zeros of $\tilde{q}$ (and hence poles of ψ) approach the unit circle, and it is in this case also that the truncation approach will run into difficulties, since the c_j will then tend to zero relatively slowly.

In tests of the program with values of r up to 20 the run time was of the order of $0.8r^3$ milliseconds on an ICL 2976 computer.

6. OTHER ALGORITHMS

In addition to simple truncation, two other methods for the singular-value analysis of infinite Hankel matrices have been proposed. We can arrive at that of Delsarte, Genin, and Kamp [4] by following the strategy outlined above but choosing a different basis for the space K_q of rational functions. Suppose that

$$q(z)=(1-\alpha_1 z)(1-\alpha_2 z)\cdots(1-\alpha_r z)$$

and that $\alpha_1,\dots,\alpha_r$ are distinct. Then $a_1,\dots,a_r$ is a basis of K_q, where $a_j(z)=(1-\alpha_j z)^{-1}$. This basis shares with the one described in Section 3 two useful properties: it yields a simple matrix for the operator $S|K_q$ (to wit, $D=\operatorname{diag}\{\alpha_1,\dots,\alpha_n\}$), and it has an easily obtained Gram matrix G_a. We have $G_a=[\gamma_{ij}]$, where

$$\gamma_{ij}=(a_j,a_i)=(1-\alpha_j\bar{\alpha}_i)^{-1}.$$

As in steps 7 and 8 of the above algorithm, we can find Q_a such that $Q_aQ_a^*=G_a^{-1}$, whereupon the desired singular values are those of $A=Q_a^{-1}h(D)Q_a$. Now the singular-value equation

$$(\lambda I-A^*A)x=0$$

reduces to

$$(\lambda G_a-h(D)^*G_a h(D))Q_a x=0,$$

that is,

$$\left[\frac{\lambda-w_j\bar{w}_i}{1-\alpha_j\bar{\alpha}_i}\right]Q_a x=0,$$

where $w_j = h(\alpha_j)$ [note that $h(D) = \mathrm{diag}\{w_1, \ldots, w_r\}$].

The major disadvantage here is that one has to know the α_j (the poles of ψ). One could of course solve for them, but as we have seen, such a step can be avoided. It could happen that we do know the α's: this will be so, for example, if the matrix H is derived from a Nevanlinna-Pick interpolation problem at a finite number of known points. Here a second objection could make itself felt: the requirement that α's be distinct. This suggests numerical instability in the case of very close α's. For this reason I believe that, when the α's are known, it is better to use a different basis, which is orthonormal and does not require distinct α's. Dr. A. C. Allison and I have in fact written a program to solve the Nevanlinna-Pick problem using this basis, and it has proved very satisfactory. Minor modifications of this program would enable it to be used to solve the singular-value problem for H.

The method of Kung [10] actually produces the eigenvalues of H rather than its singular values; however, if the entries of H are real, then H is symmetric and the two notions coincide. In terms of our notation Kung proves the following.

Let the entries of H be real, and let $\tilde{q}(z) = z^r q(1/z)$, where q is given by (5). Let

$$\frac{\tilde{q}(z)}{q(z)} = f_0 + f_1 z + f_2 z^2 + \cdots .$$

Then the singular values of H are the solutions λ of the generalized eigenvalue problem

$$\begin{bmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_1 & c_2 & \cdots & c_n \\ \vdots & \vdots & & \vdots \\ c_{n-1} & c_n & \cdots & c_{2n-1} \end{bmatrix} X = \lambda \begin{bmatrix} f_n & f_{n-1} & \cdots & f_1 \\ f_{n+1} & f_n & \cdots & f_2 \\ \vdots & \vdots & & \vdots \\ f_{2n-1} & f_{2n-2} & \cdots & f_n \end{bmatrix} X,$$

and the singular vector of H corresponding to λ is the sequence whose generating function is $(x_0 + x_1 z + \cdots + x_{n-1} z^{n-1})/q(z)$, when $x = (x_0, \ldots, x_{n-1})^T$.

It looks as though Kung's algorithm should be of comparable efficiency to the one described here, for real H. Two points appear to be slightly in favor of the present one: firstly, ours reduces to a straight singular-value problem, whereas Kung's reduces to a generalized eigenvalue problem, and algorithms for the former are somewhat faster and more reliable. Secondly, the calculation of the f_j's will undoubtedly introduce more rounding error than that of the entires of G^{-1} [see equation (13)]. However, it will require practical tests to decide between the two methods.

7. MATRICES OF INFINITE OR UNKNOWN RANK

Suppose we wish to calculate the largest singular value (i.e. operator norm) of an infinite Hankel matrix $H = [c_{i+j}]$ which does not have known finite rank. We can simply truncate after some large number of rows and columns: this will give us a lower bound for the desired value. Another idea would be to try and use the present algorithm. We could supply as data some large integer r and the first $2r$ entries of the first row of H. The program will then compute the singular values of the unique infinite Hankel matrix of rank r having the first $2r$ entries in its first row equal to those of H: let us call this the r-extrapolation of H. The r-extrapolation may or may not be a better approximation to H than the simple truncation. If we are very lucky and the rank of H was in fact no greater than r, then the r-extrapolation of H is exactly H. If we are unlucky, the r-extrapolation may not even be bounded: suppose $c_0 = c_1 = c_2 = c_3 = 1$ and $c_j = 0$ for $j > 3$. In this case the 2-extrapolation has all its entries equal to 1, and does not represent a bounded operator. This would be detected in the algorithm at step 7, the Choleski decomposition of G^{-1}. If H is unbounded, then q will have a zero in the unit disc and G^{-1} will not be positive definite. This occurs if one takes H to be the Hilbert matrix $[c_k = 1/(k+1)]$ and r to be sufficiently large, though it may be caused by rounding errors.

In spite of this difficulty I think this idea, with modifications, worth trying as a method of tackling the identification problem [3]. Suppose we know the impulse response of a system—that is, we know c_j for $j = 0$ up to some large number N, and we wish to approximate $\sum_0^\infty c_j z^j$ by a rational function using the Hankel norm criterion (see [6]). The degree r of the denominator of this rational approximant will be limited by practical considerations. Attempting to match the first $N+1$ coefficients by a rational function of degree $r-1$ in the numerator and r in the denominator leads to an overdetermined system of linear equations for large N. We can take a least-squares (or weighted least-squares) solution to obtain a rational function ψ. Passing from ψ to an equivalent minimal-norm approximant requires the singular-value analysis of the Hankel operator T_ψ, which can be carried out by the above method.

The referee has pointed out the relevance of [11] to this question.

REFERENCES

1 V. M. Adamyan, D. Z. Arov, and M. G. Krein, Infinite Hankel matrices and generalized Caratheodory-Fejér and Riesz problems, *Funkcional Anal. i Priložen.* 2:1–19 (1968).

2 R. P. Brent, F. G. Gustavson, and D. Y. Yun, Fast solution of Toeplitz systems of equations and computation of Padé approximants, *J. Algorithms* 1:219–295 (1980).

3 A. Bultheel and P. Dewilde, On the Adamyan-Arov-Krein approximation, identification and balanced realization of a system, presented at ECCTD Conference, Warsaw, 1980.
4 Ph. Delsarte, Y. Genin, and Y. Kamp, On the role of the Nevanlinna-Pick problem in circuit and system theory, *Circuit Theory Appl.* 9:177–187 (1981).
5 W. L. Duren, *Theory of H^p Spaces*, Academic, 1970.
6 Y. Genin and S. Y. Kung, A two-variable approach to the model reduction problem with Hankel norm criterion, to appear.
7 M. H. Gutknecht and L. M. Trefethen, Recursive digital filter design by the Caratheodory-Fejér method, Internal Report, Dept. of Computing Science, Stanford Univ.
8 J. W. Helton, The distance of a function to H^∞ in the Poincaré metric: electrical power transfer, *J. Funct. Anal.* 38:273–314 (1980).
9 K. Hoffman, *Banach Spaces of Analytic Functions*, Prentice-Hall, Englewood Cliffs, N.J., 1962.
10 S.-Y. Kung, Optimal Hankel-norm model reductions—scalar systems, in *Proceedings of the JACC*, San Francisco, 1980.
11 S.-Y. Kung, A new identification and model reduction algorithm via singular value decomposition, in *Proceedings of the 12th Asilomar Conference on Circuits, Systems and Computers*, Nov. 1978, pp. 705–714.
12 D. Sarason, Generalized interpolation in H^∞, *Trans. Amer. Math. Soc.* 127:179–203 (1967).
13 R. R. Whitehead, Moment methods and Lanczos methods, in *Theory and Applications of Moment Methods in Many Fermion Systems*, (B. J. Dalton, S. M. Grimes, J. P. Vany and S. A. Williams, Eds.), Plenum, 1980.
14 N. J. Young, The solution of Lyapunov's equation by a geometric method, *Proc. Roy. Soc. Edinburgh* 86A:347–354 (1980).
15 N. J. Young, An identity which implies Cohn's theorem on the zeros of a polynomial, *J. Math. Anal. Appl.* 70:240–248 (1979).

Received 24 December 1981; revised 27 September 1982