COMPUTATION OF A UNIMODULAR MATRIX

by

A. Kontos and J. B. Pearson

December 1979

Technical Report No. 7916

COMPUTATION OF A UNIMODULAR MATRIX

by

A. Kontos* and J. B. Pearson*

Abstract

An algorithm for computing a unimodular matrix $U(\lambda)$ satisfying the equation $[A(\lambda)\ B(\lambda)]\ U(\lambda) = [I\ 0]$ is presented where $A(\lambda)$ and $B(\lambda)$ are relatively left prime polynomial matrices. The approach used avoids Euclidean-type operations used in standard Gaussian elimination and thus appears to be superior to any direct approach based on the computation of Smith forms.

---

* Department Of Electrical Engineering, Rice University, Houston, TX. 77001. This research was supported by the National Science Foundation under Grant ENG 77-04119.

1. Introduction

Given the left coprime polynomial matrices $A(\lambda)$ and $B(\lambda)$ of order $n \times n$ and $n \times m$ respectively, it is well known [pg.14,33] that there exists a unimodular matrix $U(\lambda)$ such that

$$[A(\lambda)\ B(\lambda)]U(\lambda) = [I_n\ 0] \tag{1.1}$$

The form $[I_n\ 0]$ is obtained by a sequence of elementary column operations on $[A(\lambda)\ B(\lambda)]$, i.e., Gaussian elimination without pivoting, which is undesirable for numerical computation.

Here a new method, composed of two steps, is proposed to find the matrix $U(\lambda)$.

(a) Given $P(\lambda) = [A(\lambda)\ B(\lambda)]$, with linearly independent rows for any λ , a matrix $Q(\lambda)$ is determined, using state feedback techniques, such that

$$R(\lambda) = \begin{bmatrix} P(\lambda) \\ Q(\lambda) \end{bmatrix}$$

is unimodular, and

(b) A block Levinson recursion method [4] [12] [13] is then used to find the inverse of $R(\lambda)$,

$$U(\lambda) = R^{-1}(\lambda)$$

which is the solution of (1.1)

2. Computation of $Q(\lambda)$

Given the nxm $(n \leq m)$ polynomial matrix $P(\lambda)$ with linearly independent rows for every λ, we will show how to find a polynomial matrix $Q(\lambda)$ such that the matrix

$$R(\lambda) = \begin{bmatrix} P(\lambda) \\ Q(\lambda) \end{bmatrix}$$

is unimodular.

Let

$$P(\lambda) = P_o + \lambda P_1 + \ldots + \lambda^{\ell} P_{\ell}$$

be the given polynomial matrix $P(\lambda)$.

Since $P(\lambda)$ has linearly independent rows for every λ, $P_o(= P(0))$ has linearly independent rows. Therefore, there exist constant non-singular matrices T_1 and T_1 such that

$$T_1 P_o T_2 = [I_n \ 0] .$$

Then

$$\overline{P}(\lambda) = T_1 P(\lambda) T_2$$

has linearly independent rows for every λ and if

$$\overline{R}(\lambda) = \begin{bmatrix} \overline{P}(\lambda) \\ \overline{Q}(\lambda) \end{bmatrix}$$

is unimodular, then

$$R(\lambda) = \begin{bmatrix} P(\lambda) \\ \overline{Q}(\lambda) T_2^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} T_1 & 0 \\ 0 & I_{n-m} \end{bmatrix}^{-1} \overline{R}(\lambda) T_2^{-1}$$

is also unimodular.

Hence there is no loss of generality to assume that $P(\lambda)$ has the form

$$P(\lambda) = [I_n \ 0] + \lambda P_1 + \ldots + \lambda^{l} P_{l} . \qquad (2.1)$$

The following Lemma is now in order.

Lemma 2.1. Let $P(\lambda)$ be given by (2.1) with linearly independent rows for every λ. There exists an $(m-n)\times m$ polynomial matrix

$$Q(\lambda) = [0 \ I_{m-n}] + \lambda Q_1 + \ldots + \lambda^q Q_q \qquad (2.2)$$

such that the matrix

$$R(\lambda) = \begin{bmatrix} P(\lambda) \\ Q(\lambda) \end{bmatrix} \qquad (2.3)$$

is unimodular.

Proof. The existence of a polynomial matrix $Q(\lambda)$ such that $R(\lambda)$ is unimodular is shown in [8, pg.71].

Let

$$\overline{Q}(\lambda) = \overline{Q}_0 + \lambda \overline{Q}_1 + \ldots + \lambda^q \overline{Q}_q$$

be such that

$$\overline{R} = \begin{bmatrix} P \\ \overline{Q} \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ \overline{Q}_{01} & \overline{Q}_{02} \end{bmatrix} + \lambda \overline{R}_1 + \ldots$$

is unimodular.

That implies that $\overline{Q}_{02}$ is invertible and

$$R(\lambda) = \begin{bmatrix} I_n & 0 \\ 0 & \overline{Q}_{02}^{-1} \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -\overline{Q}_{01} & I_{m-n} \end{bmatrix} \overline{R}(\lambda)$$

$$= \begin{bmatrix} I_n & 0 \\ 0 & I_{m-n} \end{bmatrix} + \lambda \begin{bmatrix} P_1 \\ Q_1 \end{bmatrix} + \ldots$$

is unimodular and the proof is complete.

In the remainder of this section we construct such a $Q(\lambda)$.

The matrix $R(1/\lambda)$ is a proper rational matrix and has the obvious controllable representation [9]

$$\tilde{S} = (\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}) \quad \text{with}$$

$$\tilde{A} = \begin{pmatrix} 0 & I_m & 0 & & 0 \\ & & & & \\ & & & & I_m \\ 0 & & & & 0 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ I_m \end{pmatrix}$$

$$\tilde{C} = \begin{pmatrix} 0 . . .0 & P_t . . .P_1 \\ & \\ Q_q & . \quad . \quad . \quad Q_1 \end{pmatrix} \quad \text{if } q > t \tag{2.4}$$

or

$$\tilde{C} = \begin{pmatrix} P_t & . \quad . \quad . \quad P_1 \\ & \\ 0 . . . 0 & Q_q . . Q_1 \end{pmatrix} \quad \text{if } q \leq t$$

and

$$\tilde{D} = I_m$$

Since $\tilde{D} = I_m$, $\tilde{S}$ has an inverse which we call

$$\hat{S} = (\hat{A}, \hat{B}, \hat{C}, \hat{D}) = (\tilde{A}-\tilde{B}\tilde{C}, \tilde{B}, -\tilde{C}, \tilde{D}) \tag{2.5}$$

An useful characterization of a unimodular matrix is the following.

Lemma 2.2. *All the eigenvalues of* $\hat{A}$ *are equal to zero* (*i.e.*, $\hat{A}$ *is nilpotent*), *if and only if* $R(\lambda)$ *is unimodular.*

Proof. (if) The transfer function of $\hat{S}$ is $R^{-1}(\frac{1}{\lambda})$ and, since $R(\lambda)$ is unimodular, $R^{-1}(\frac{1}{\lambda})$ is a polynomial in $1/\lambda$.

The realization (A,B,C,D) is a controllable one and the fact that $R^{-1}(\frac{1}{\lambda})$ has all its poles at zero shows that the observable modes of the pair $(\hat{A},\hat{C})$ correspond to zero eigenvalues of $\hat{A}$.

The unobservable modes correspond to eigenvalues λ that reduce the

$$\text{rank}\begin{pmatrix}\lambda I - \hat{A}\\ \hat{C}\end{pmatrix} = \text{rank}\begin{pmatrix}\lambda I - \tilde{A}\\ \tilde{C}\end{pmatrix}$$

Therefore, if there are any unobservable modes they correspond to eigenvalues of $\tilde{A}$ which are all zero.

(only if). Let A be nilpotent. There exists a positive finite integer ζ such that $\hat{A}^{\zeta} = 0$, and $\hat{A}^{k} = 0$ for $k \geq \zeta$. It follows that

$$R^{-1}\left(\frac{1}{\lambda}\right) = \hat{D} + \hat{C}\,(\lambda I - \hat{A})^{-1}\hat{B}$$

or

$$\begin{aligned} R^{-1}(\lambda) &= \hat{D} + \hat{C}\,\left(\frac{1}{\lambda}I - \hat{A}\right)^{-1}\hat{B} \\ &= \hat{D} + \sum_{k=1}^{\infty} \lambda^{k}\hat{C}\hat{A}^{k-1}\hat{B} \\ &= \hat{D} + \sum_{k=1}^{\zeta} \lambda^{k}\hat{C}\hat{A}^{k-1}\hat{B} \end{aligned}$$

which is a polynomial.

Therefore $R(\lambda)$ is unimodular.

Q.E.D.

Now define

$$A = \begin{pmatrix} 0 & I_m & & 0 \\ & & & \\ & & & I_m \\ \begin{pmatrix}-P_{\iota}\\ 0\end{pmatrix} & \cdot & \cdot & \cdot\begin{pmatrix}-P_1\\ 0\end{pmatrix} \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ I_{m-n} \end{pmatrix} \qquad (2.6)$$

Lemma 2.3. Uncontrollable modes of the pair (A,B) *correspond to zero eigenvalues of* A .

Proof. From (2.4) and (2.5) we have

$$\hat{A} = \left[\begin{array}{cc|cc} 0 & 1 & & \\ & & 1 & 0 \\ & & 0 & 1 \quad 0 \\ \hline & 0 & & A \end{array}\right] + \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ \hline B \end{array}\right] [-Q_q \cdots -Q_1] \qquad \text{if } q > \ell$$

or

$$\hat{A} = A + B[0 \cdots 0 - Q_q \cdots -Q_1] \qquad \text{if } q \leq \ell$$

In each case eigenvalues of A corresponding to uncontrollable modes of (A,B) must be eigenvalues of $\hat{A}$ which by assumption are zero.

Q.E.D.

The main result follows immediately.

Theorem 2.4. Let A *and* B *be defined in* (2.6) . *There exists a matrix*

$$F = [\underbrace{-Q_\ell}_{\leftarrow m \rightarrow} \cdot \quad \cdot \quad \cdot \underbrace{-Q_1}_{\leftarrow m \rightarrow}] \tag{2.7}$$

such that the matrix A+BF *is nilpotent. Furthermore, if* $R(\lambda)$ *is defined through* (2.3), (2.2) *and* (2.7), *then* $R(\lambda)$ *is unimodular.*

Proof. Follows from Lemmas 2.2 and 2.3 .

Notice that the pair (A,B) is controllable if and only if P_ℓ (as defined in 2.1) has full row rank. If it is not controllable we can find a non-singular matrix T such that

$$TAT^{-1} = \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix}$$

and

and

$$TB = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}$$

and (A_1, B_1) is a controllable pair.

After that we can use a pole placement algorithm to find F as in theorem 2.4 . This is a straightforward procedure that takes no advantage of the special structure of A .

In the rest of this section we will show how to take advantage of the fact that A already has some eigenvalues equal to zero.

With A and B as defined in (2.6), define the system

$$S : \dot{x} = Ax + Bu \tag{2.8}$$

Define the sequence of systems S_i as follows

$$S_0 : \dot{z}_0 = A_0 z_0 + B_0 u \tag{2.9a}$$

where

$$A_0 \triangleq A , B_0 \triangleq B , z_0 \triangleq x \tag{2.9b}$$

The rest of the sequence is defined inductively. For the induction step, let

$$S_{k-1} : \dot{z}_{k-1} = A_{k-1} z_{k-1} + B_{k-1} u \tag{2.9c}$$

A_{k-1} is an $r_{k-1} x r_{k-1}$ matrix of rank ρ_{k-1} . If $r_{k-1} = \rho_{k-1}$, define $a = k-1$ and S_a is the last member of the sequence (it may be $r_a = 0$) .

If $r_{k-1} > \rho_{k-1}$, then there exist matrices L_k and M_k of order $r_{k-1} x \rho_{k-1}$ and $\rho_{k-1} x r_{k-1}$, respectively, and rank ρ_{k-1} such that

$$A_{k-1} = L_k M_k \quad . \tag{2.9d}$$

Define

$$z_k = M_k z_{k-1} \tag{2.9e}$$

then

$$\dot{z}_k = M_k L_k M_k z_{k-1} + M_k B_{k-1} u$$

$$= M_k L_k z_k + M_k B_{k-1} u$$

or

$$S_k : \dot{z}_k = A_k z_k + B_k u \tag{2.9f}$$

where

$$A_k \triangleq M_k L_k \quad , \quad B_k = M_k B_{k-1} \, . \tag{2.9g}$$

Notice that $r_k = \rho_{k-1} < r_{k-1}$ and since we start with a finite order system the process described above is going to terminate in a finite number of steps.

Lemma 2.5. If $r_a > 0$ the pair (A_a, B_a) is controllable.

Proof. It is easy to see that

$$z_a = M_a M_{a-1} \cdots M_1 x \triangleq Mx \tag{2.10a}$$

$$\begin{aligned} A_a z_a &= M_a (L_a M_a) M_{a-1} \cdots M_1 x \\ &= M_a (M_{a-1} L_{a-1}) M_{a-1} \cdots M_1 x \\ &= M_a \cdots M_1 L_1 M_1 x \\ &= MAx \end{aligned} \tag{2.10b}$$

and

$$B_a = MB \tag{2.10c}$$

where

$$M = M_a M_{a-1} \cdots M_1 \tag{2.11}$$

If $x^T M = 0$ for some vector x, then

$$x^T M_a \cdots M_2 = 0 \quad , \text{ since } M_1 \text{ has full row rank,}$$

or

$$x^T M_a \cdots M_3 = 0$$

or

$$x = 0$$

that is, M has full row rank. Let

$$W = \begin{bmatrix} M \\ N \end{bmatrix}$$

be square and non-singular for some matrix N .

The transformation

$$\begin{bmatrix} z_a \\ y \end{bmatrix} = Wx$$

transforms the system S , according to (2.10) , to the form

$$\overline{S} : \begin{bmatrix} \dot{z}_a \\ \dot{y} \end{bmatrix} = \begin{bmatrix} A_a & 0 \\ H_2 & H_1 \end{bmatrix} \begin{bmatrix} z_a \\ y \end{bmatrix} + \begin{bmatrix} B_a \\ R \end{bmatrix} u$$

for some H_1, H_2 and R .

From Lemma 2.3 we know that the uncontrollable modes of S correspond to zero eigenvalues and the same is true for $\overline{S}$, which implies that the matrix

$$\begin{bmatrix} \lambda I-A_a & 0 & B_a \\ -H_2 & \lambda I-H_1 & R \end{bmatrix}$$

has full rank for every $\lambda \neq 0$, and

$$[\lambda I-A_a \;\; B_a]$$

has full rank for every $\lambda \neq 0$. Since A_a is non-singular we conclude that

$$[\, -A_a \;\; B_a \,]$$

has full rank also,

i.e.

$$[\lambda I - A_a \ \ B_a]$$

has full rank for every λ .

Q.E.D.

Lemma 2.6.

$$|\lambda I - (A_k + B_k F_k)| = \lambda^{d_{k+1}} |\lambda I - (A_{k+1} + B_{k+1} F_{k+1})| \qquad (2.12)$$

where

$$k = 0,1,2,\cdots$$

$$F_k = F_{k+1} M_{k+1} \quad \text{and} \quad d_{k+1} = r_{k+1} - r_{k+2} \ .$$

Proof. If follows that

$$\begin{aligned} |\lambda I - (A_k + B_k F_k)| &= |\lambda I - (L_{k+1} M_{k+1} + B_k F_{k+1} M_{k+1})| \\ &= \lambda^{r_{k+1} - \rho_{k+1}} |\lambda I - M_{k+1}(L_{k+1} + B_k F_{k+1})| * \\ &= \lambda^{d_{k+1}} |\lambda I - (A_{k+1} + B_{k+1} F_{k+1})| \ . \end{aligned}$$

We are now ready to present the following.

Theorem 2.7. Let A, B be defined by (2.6) , A_a, B_a be defined by (2.9) and M be defined by (2.11)

Then

(i) If $r_a > 0$ there exists a matrix F_a such that $(A_a + B_a F_a)$ is nilpotent, and with $F = F_a M$, $A + BF$ is nilpotent.

(ii) If $r_a = 0$, then A is nilpotent.

Proof.

(i) The existence of Fa is implied directly by the controllability of the pair (A_a, B_a) .

---

* This is because

$$|\lambda I - LM| = \lambda_{m-n} |\lambda I - ML|$$

where L and M are mxn and nxm (m>n) constant matrices, respectively.

Applying (2.12) for $k = 0,1,\cdots,a$, we have

$$|\lambda I - (A + BF_0)| = \lambda^{d_1}|\lambda I - (A_1 + B_1F_1)|$$

$$|\lambda I - (A_1 + B_1F_1)| = \lambda^{d_2}|\lambda I - (A_2 + B_2F_2)|$$

$$\vdots$$

$$|\lambda I - (A_{a-1} + B_{a-1}F_{a-1})| = \lambda^{d_a}|\lambda I - (A_a + B_aF_a)|$$

Multiplying all the above equalities, we get

$$|\lambda I - (A + BF)| = \lambda^{d_1+\cdots+d_a}|\lambda I - (A_a + B_aF_a)| \tag{2.13}$$

where

$$\begin{aligned} F \overset{\Delta}{=} F_0 &= F_1M_1 \\ &= F_2M_2M_1 \\ &= \cdots \\ &= F_aM \end{aligned}$$

Now choose F_a such that

$$|\lambda I - (A_a + B_aF_a)| = \lambda^{r_a} .$$

Then from (2.13) we have

$$|\lambda I - (A + BF)| = \lambda^{r} .$$

(ii) If $r_a = 0$, the same argument as above shows that A is nilpotent.

Q.E.D.

In this section we have shown how, given $P(\lambda)$ as in (2.2) we can find $Q(\lambda)$ such that $R(\lambda) = \begin{bmatrix} P(\lambda) \\ Q(\lambda) \end{bmatrix}$ is unimodular. The method is illustrated with the following example.

<u>Example</u>

Let $$P(\lambda) = [1 + \alpha\lambda^2 \;\; \beta\lambda^2]$$

where α and β are nonzero real numbers. Using the procedure given above

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\alpha & -\beta & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Since $r_o = 4$, $\rho_o = 3$, it follows that

$$A_o = L_1 M_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\alpha & -\beta & 0 & 0 \end{bmatrix}$$

$$A_1 = M_1 L_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -\alpha & -\beta & 0 \end{bmatrix} \qquad B_1 = M_1 B_o = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Here $r_1 = 3$, $\rho_1 = 2$ so we continue

$$A_1 = L_2 M_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ -\alpha & -\beta & 0 \end{bmatrix}$$

$$A_2 = M_2 L_2 = \begin{bmatrix} 0 & 1 \\ -\alpha & 0 \end{bmatrix} \qquad B_2 = M_2 B_1 = \begin{bmatrix} 0 \\ -\beta \end{bmatrix}$$

Since $r_2 = 2 = \rho_2$, $a = 2$ and (A_2, B_2) is controllable.

With $F_2 = [-\frac{\alpha}{\beta} \ 0]$ $(A_2 + BF_2)$ is nilpotent and

$$M = M_2 M_1 = \begin{bmatrix} -\alpha & -\beta & 0 & 0 \\ 0 & 0 & -\alpha & -\beta \end{bmatrix}$$

$$F = F_2 M = \begin{bmatrix} \frac{\alpha^2}{\beta} & \alpha & 0 & 0 \end{bmatrix}$$

Therefore

$$Q(\lambda) = [\ 0 \ 1\ ] - [\ 0 \ 0\ ]\lambda - [\ \frac{\alpha^2}{\beta} \ \alpha\ ]\lambda^2$$

and

$$R(\lambda) = \begin{bmatrix} 1 + \alpha\lambda^2 & \beta\lambda^2 \\ -\frac{\alpha^2}{\beta}\lambda^2 & 1 - \alpha\lambda^2 \end{bmatrix}$$

is unimodular.

<u>Remark</u>

In the previous example M_1 can be obtained by deleting the zero rows of A_o, and then A_1 can be obtained by deleting the corresponding columns. M_2 and A_2 can be obtained from A_1 by deleting the zero rows and columns, and so on. This observation can be used to save some computation time and storage. More specifically, these trivial operations can be performed at least $\iota(m-n)$ times, and result to an A matrix of order at the most $\iota n \times \iota n$.

3. Inverse of a Unimodular Matrix

For any $n x n$ unimodular matrix

$$R(\lambda) = I_n + \lambda R_1 + \cdots + \lambda^{\iota} R_{\iota} \tag{3.1}$$

there is an inverse $U(\lambda) = R^{-1}(\lambda)$ of the form

$$U(\lambda) = I_n - \lambda U_1 - \cdots - \lambda^d U_d \tag{3.2}$$

for some finite integer d. Obviously

$$R(\lambda)U(\lambda) = I_n \tag{3.3}$$

or, by equating the coefficients of the powers of λ

$$\begin{bmatrix} I_n & & & \\ R_1 & I_n & \ddots & \\ \cdot & R_1 & & I_n \\ \cdot & \vdots & & R_1 \\ R_{\iota} & \cdot & & \cdot \\ & R_{\iota} & & \cdot \\ & & & R_{\iota} \end{bmatrix} \begin{bmatrix} I_n \\ -U_1 \\ \\ \\ -U_d \end{bmatrix} = \begin{bmatrix} I_n \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

It should be mentioned here that d cannot be determined from ι and n only, as is shown in

Example:

(a)

$$R_1(\lambda) = I_3 + \lambda \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$U_1(\lambda) = R_1^{-1}(\lambda) = I_3 + \lambda \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\iota = 1, \quad n = 3, \quad d = 2$

(b)

$$R_2(\lambda) = I_3 + \lambda \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$U_2(\lambda) = R_2^{-1}(\lambda) = I_3 + \lambda \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\iota = 1, \quad n = 3, \quad d = 1$.

In both cases we have $\iota = 1$ and $n = 3$ but d is different in the two cases. The next Lemma provides an upper bound on d .

Lemma 3.1. $d \leq \iota n - 1$

Proof.

From the proof of Lemma 2.2 it follows that

$$R^{-1}(\lambda) = U(\lambda) = \hat{D} + \sum_{k=1}^{\alpha} \lambda^k \hat{C}\hat{A}^{k-1}\hat{B} \tag{3.5}$$

for some positive integer α such that $\hat{A}^{\alpha} = 0$. Since $\hat{A}$ is $(\iota n) \times (\iota n)$, we conclude that $\alpha \leq \iota n$. Let $d = \alpha - 1$, then

$$\hat{A}^{\alpha} = \hat{A}\,\hat{A}^d = 0$$

and from (2.4) and (2.5)

$$\hat{A} = \begin{bmatrix} X \\ \hat{C} \end{bmatrix} \quad \text{for some } X .$$

Therefore $\hat{C}\,\hat{A}^d = 0$
and

$$U(\lambda) = \hat{D} + \sum_{k=1}^{d} \lambda^k \hat{C}\hat{A}^{k-1}\hat{B}$$

with $d \leq \iota n - 1$

Q.E.D.

As is shown in the first example preceeding this Lemma, the bound given by Lemma 3.1 is tight $(d = \ell n - 1)$.

We now return to the equation (3.4) . We will solve (3.4) by induction on d , i.e., we will assume that $d = 1$ (denote this U_1 as $U_1^{(1)}$ to show that is the first element of a length one solution) and we will minimize some norm of

$$E_1 = \begin{pmatrix} R_1 \\ \cdot \\ \cdot \\ R_\ell \\ 0 \\ \cdot \\ \cdot \end{pmatrix} - \begin{pmatrix} I_n \\ R_1 \\ \cdot \\ \cdot \\ R_\ell \\ \cdot \\ \cdot \end{pmatrix} \begin{pmatrix} U_1^{(1)} \end{pmatrix}$$

If the norm of E_1 is not small enough, using $U_1^{(1)}$, we will find a length two solution

$$U^{(2)} \triangleq \begin{pmatrix} U_1^{(2)} \\ U_2^{(2)} \end{pmatrix}$$

and so on until for some length k

$$U^{(k)} = \begin{pmatrix} U_1^{(k)} \\ \vdots \\ U_k^{(k)} \end{pmatrix}$$

a suitably small norm can be obtained for

$$E_k = \begin{pmatrix} R_1 \\ R_2 \\ \cdot \\ \cdot \\ R_\ell \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} - \begin{pmatrix} I_n & & \\ R_1 & I_n & \\ & R_1 & \\ & \cdot & I_n \\ & \cdot & R_1 \\ R_\ell & \cdot & \cdot \\ 0 & R_\ell & \cdot \\ \cdot & & \cdot \\ \cdot & & R_\ell \end{pmatrix} \begin{pmatrix} U_1^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ U_k^{(k)} \end{pmatrix} \qquad (3.6)$$

Notice that for $k=d$ and $E_k=0$ (3.6) is equivalent to (3.4) .

We now focus our attention on (3.6) which we write as

$$\left[\epsilon_1^k \cdots \epsilon_n^k\right] = \left[b_1 \dots b_n\right] - A_k^k \left[u_1^k \dots u_n^k\right] \qquad (3.7)$$

where

$$\left[\epsilon_1^k \cdots \epsilon_n^k\right] = E_k$$

$$\left[b_1 \dots b_n\right] = B \triangleq \begin{bmatrix} R_1 \\ R_2 \\ \cdot \\ \cdot \\ R_\ell \\ 0 \\ \cdot \\ \cdot \end{bmatrix}$$

$$\left[u_1^k \dots u_n^k\right] = U^{(k)} = \begin{bmatrix} U_1^{(k)} \\ \vdots \\ U_k^{(k)} \end{bmatrix}$$

and

$$A_k = \underbrace{\begin{bmatrix} I_n & \\ R_1 & \\ \cdot & I_n \\ \cdot & R_1 \\ \cdot & \cdot \\ R_\ell & \cdot \\ 0 & \cdot \\ \cdot & R_\ell \end{bmatrix}}_{\leftarrow k \text{ blocks} \rightarrow}$$

clearly A_k has full column rank. The equation (3.7) is equivalent to the system of equations

$$\epsilon_i^k = b_i - A_k u_k^k \qquad i=1,2,\dots,n \qquad (3.8)$$

We will find u i=1,2,...,n so that $\|\epsilon_i^k\|_2$ is minimized $(\|\epsilon\|_2 = \sqrt{\epsilon^T\epsilon})$.

This is equivalent to minimizing $\|\epsilon_1^k\|_2^2 + \|\epsilon_2^k\|_2^2 + \cdots + \|\epsilon_n^k\|_2^2$ since the ϵ_i^k i=1,2,...,n are independent. With $\|E_k\|_T^2 = \|\epsilon_1^k\|_2^2 + \cdots + \|\epsilon_n^k\|_2^2$, $\|E_k\|_F$ is the well known norm.

It is well known (see for example [10]) that $\|\epsilon_i^k\|_2$ is minimized if

$$A_k^T \epsilon_i^k = 0 \qquad (3.9)$$

Applying (3.9) for i=1,2,...,n we obtain

$$A_k^T E_k = 0 \qquad (3.10)$$

which implies

$$A_k^T B = A_k^T A_k U^{(k)} \qquad (3.11)$$

Define the non-singular matrix

$$L_k = A_k^T A_k$$

Also define

$$P_j = \sum_{i=0}^{i-j} R_i^T R_{i+j} \qquad j=0,1,2,... \qquad (3.12)$$

with

$$R_0 = I_n$$

It follows that

$$L_k = \begin{pmatrix} P_0 & P_1^T & \cdot & \cdot & P_{k-1}^T \\ P_1 & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & P_1^T \\ P_{k-1} & \cdot & \cdot & P_1 & P_0 \end{pmatrix} \qquad (3.13)$$

and (3.11) is equivalent to

$$L_k U^{(k)} = \begin{pmatrix} P_1 \\ \cdot \\ \cdot \\ \cdot \\ P_k \end{pmatrix} , \quad k=1,2,... \qquad (3.14)$$

Define $G^{(k)}$ through

$$L_k G^{(k)} = \begin{bmatrix} P_k^T \\ \cdot \\ \cdot \\ \cdot \\ P_1^T \end{bmatrix} , \quad k=1,2,... \qquad (3.15)$$

and write $G^{(k)}$ as

$$G^{(k)} = \begin{bmatrix} G_k^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ G_1^{(k)} \end{bmatrix} , \quad k=1,2,... \qquad (3.16)$$

We will now derive recursive formulae for the sequence $(U^{(k)},G^{(k)})$ and the error

$$e_k = E_k^T E_k \qquad (3.17)$$

using a block Levinson recursion method [4] [12] [13] .

<u>Solution for $U^{(k)}$ and $G^{(k)}$</u>

From (3.14) , (3.15) ,and (3.13)

$$P_o U_1^{(1)} = P_1$$

and

$$P_o G_1^{(1)} = P_1^T .$$

Now, given $U^{(k-1)}$ and $G^{(k-1)}$, we will find $U^{(k)}$ and $G^{(k)}$ without inverting L_k .

From (3.14)

$$\left[\begin{array}{ccc|c} & & & P_{k-1}^T \\ & & & \cdot \\ & L_{k-1} & & \cdot \\ & & & \cdot \\ & & & P_1^T \\ \hline P_{k-1} & \cdot \quad \cdot & P_1 & P_o \end{array}\right] \left[\begin{array}{c} U_1^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ U_{k-1}^{(k)} \\ \hline U_k^{(k)} \end{array}\right] = \left[\begin{array}{c} P_1 \\ \cdot \\ \cdot \\ \cdot \\ P_{k-1} \\ \hline P_k \end{array}\right]$$

which implies

$$L_{k-1}\begin{bmatrix} U_1^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ U_{k-1}^{(k)} \end{bmatrix} + \begin{bmatrix} P_{k-1}^T \\ \cdot \\ \cdot \\ \cdot \\ P_1^T \end{bmatrix} U_k^{(k)} = \begin{bmatrix} P_1 \\ \cdot \\ \cdot \\ \cdot \\ P_{k-1} \end{bmatrix}$$

and

$$\begin{bmatrix} P_{k-1} \cdot & \cdot & \cdot P_1 \end{bmatrix} \begin{bmatrix} U_1^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ U_{k-1}^{(k)} \end{bmatrix} + P_0 U_k^{(k)} = P_k$$

Multiplying the first of these two formulas from the left by L_{k-1}^{-1} and using the second, (3.14), and (3.15), it is a simple matter to check that, if

$$D_{k-1} \triangleq P_0 - [P_{k-1} \cdot \cdot \cdot P_1]\, G^{(k-1)} \tag{3.18}$$

then

$$D_{k-1} U^{(k)} = P_k - [P_{k-1}, \ldots, P_1]\, U^{(k-1)} \tag{3.19}$$

and

$$\begin{bmatrix} U_1^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ U_{k-1}^{(k)} \end{bmatrix} = \begin{bmatrix} U_1^{(k-1)} \\ \cdot \\ \cdot \\ \cdot \\ U_{k-1}^{(k-1)} \end{bmatrix} - \begin{bmatrix} G_{k-1}^{(k-1)} \\ \cdot \\ \cdot \\ \cdot \\ G_1^{(k-1)} \end{bmatrix} U_k^{(k)} \tag{3.20}$$

Starting from (3.15) , or

$$\left[\begin{array}{c|c} P_0 & P_1^T \cdot \cdot P_{k-1}^T \\ \hline P_1 & \\ \cdot & \\ \cdot & L_{k-1} \\ P_{k-1} & \end{array}\right] \begin{bmatrix} G_k^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ G_1^{(k)} \end{bmatrix} = \begin{bmatrix} P_k^T \\ \cdot \\ \cdot \\ \cdot \\ P_1^T \end{bmatrix}$$

we can find recursive formulae for $G^{(k)}$, i.e., if

$$F_{k-1} \triangleq P_0 - \begin{pmatrix} P_1^T & \cdots & P_{k-1}^T \end{pmatrix} U^{(k-1)} \tag{3.21}$$

then

$$F_{k-1} G_k^{(k)} = P_k^T - \begin{pmatrix} P_1^T & \cdots & P_{k-1}^T \end{pmatrix} G^{(k-1)} \tag{3.22}$$

and

$$\begin{pmatrix} G_{k-1}^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ G_1^{(k)} \end{pmatrix} = \begin{pmatrix} G_{k-1}^{(k-1)} \\ \cdot \\ \cdot \\ \cdot \\ G_1^{(k-1)} \end{pmatrix} - \begin{pmatrix} U_1^{(k-1)} \\ \cdot \\ \cdot \\ \cdot \\ U_{k-1}^{(k-1)} \end{pmatrix} G_k^{(k)} \tag{3.23}$$

The equations (3.18) - (3.23) give $U^{(k)}$ and $G^{(k)}$. We need only invert the matrices D_{k-1} and F_{k-1} which are nxn and as it is shown in the following Lemma they are non-singular.

<u>Lemma 3.2</u>. <u>D_{k-1} and F_{k-1} as given by (3.18) and (3.21) respectively are both non-singular</u>.

<u>Proof</u>. Let $S^T = [P_{k-1},\ldots,P_1]$ so that

$$L_k = \begin{pmatrix} L_{k-1} & S \\ S^T & P_0 \end{pmatrix}$$

Then

$$\begin{aligned} |D_{k-1}| &= \left| P_0 - [P_{k-1},\ldots,P_1] G^{(k-1)} \right| \\ &= \left| P_0 - [P_{k-1}^{k},\ldots,P_1] L_{k-1}^{-1} [P_{k-1},\ldots,P_1]^T \right| \\ &= \left| P_0 - S^T L_{k-1}^{-1} S \right| \\ &= \left| \begin{pmatrix} I & L_{k-1}^{-1} S \\ 0 & P_0 - S^T L_{k-1}^{-1} S \end{pmatrix} \right| \end{aligned}$$

$$= \left| \begin{bmatrix} I & 0 \\ -S^T & I \end{bmatrix} \begin{bmatrix} L_{k-1}^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} L_{k-1} & S \\ S^T & P_o \end{bmatrix} \right|$$

$$= \left| L_{k-1}^{-1} \right| \left| L_k \right|$$

$$\neq 0$$

Similarly,

$$\left| F_{k-1} \right| \neq 0$$

Q.E.D.

We will now show that the right side of (3.19) is equal to the transpose of the right side of (3.22) ("Burg's Lemma", [1] [13]) , a result which can be used to simplify computation.

Lemma 3.3.

$$W_k \triangleq P_k - [P_{k-1} \ldots P_1]\, U^{(k-1)} = \left(P_k^T - [P_1^T \ldots P_{k-1}^T]\; G^{(k-1)}\right)^T \tag{3.24}$$

Proof.

$$\left(P_k^T - [P_1^T \ldots P_{k-1}^T] G^{(k-1)}\right)^T = \left(P_k^T - [P_1^T \ldots P_{k-1}^T] L_{k-1}^{-1} \begin{bmatrix} P_{k-1}^T \\ \cdot \\ \cdot \\ P_1^T \end{bmatrix} \right)^T$$

$$= P_k - [P_{k-1} \ldots P_1] L_{k-1}^{-1} \begin{bmatrix} P_1 \\ \cdot \\ \cdot \\ P_{k-1} \end{bmatrix}$$

$$= P_k - [P_{k-1} \ldots P_1]\; U^{(k-1)}$$

Q.E.D.

So far we have shown how to obtain $U^{(k)}$ and $G^{(k)}$, given $U^{(k-1)}$ and $G^{(k-1)}$ inverting two nxn matrices.

We will next derive a recursive formula for e_k as defined in (3.17) .

Solution for e_k

From (2.17)

$$
\begin{aligned}
e_k &= E_k^T E_k \\
&= \left[B - A\, U^{(k)}\right]^T E_k \\
&= B^T E_k - U^{(k)T} A_k^T E_k \\
&= B^T E_k \\
&= B^T B - B^T A_k U^{(k)} \\
&= - I_n + P_o - \left[P_1^T \ldots P_k^T\right] U^{(k)} \\
&= - I_n + P_o - \left[P_1^T \ldots P_{k-1}^T\right] \begin{bmatrix} U_1^{(k)} \\ \cdot \\ \cdot \\ U_{k-1}^{(k)} \end{bmatrix} - P_k^T U_k^{(k)} \\
&= - I_n + P_o - \left[P_1^T \ldots P_{k-1}^T\right] \left[U^{(k-1)} - G^{(k-1)} U_k^{(k)}\right] - P_k^T U_k^{(k)} \\
&= e_{k-1} - \left[P_k^T - [P_1^T \ldots P_{k-1}^T]\, G^{(k-1)}\right] U_k^{(k)} \\
&= e_{k-1} - W_k^T U_k^{(k)} \qquad \text{for } k = 1,2,\ldots \qquad (3.25)
\end{aligned}
$$

and

$$e_o = - I_n + P_o \qquad (3.25)$$

where use was made of (3.10), (3.12), (3.20) and (3.24) .

Define $\| e_k \|_T = \| E_k \|_F$. Then clearly we must have $\| e_k \|_T \geq \| e_{k-1} \|_T$ for $k=1,2,\ldots$, and since $R(\lambda)$ is unimodular, there exists $d \leq \iota n-1$ such that $\| e_d \|_T = 0$. In practice, however, due to round off errors, $\| e_k \|_T$ almost always will be non-zero. This suggests that in practice there are no unimodular matrices.

The situation is similar to the singularity of a constant square matrix. More specifically, a constant square matrix is singular if its smallest singular value is zero. Using finite precision conputations a singular matrix will usually have a non-zero smallest singular value, although very close to zero.

We say that a matrix is singular if its smallest singular value is smaller than some small positive number ϵ. This is widely acceptable and it makes sense in several applications [11] [7] [3] .

In the case of unimodular matrices we call a matrix unimodular if for some $d \le \iota n - 1$, $\| e_d \|_T$ is less than some small positive number ϵ .

The following example illustrates the method.

Example.

$$R(\lambda) = I_2 + \lambda \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \lambda^2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad P_1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad P_3 = \ldots = 0$$

$$U_1^{(1)} = \begin{pmatrix} 0 & 1 \\ 0 & \frac{1}{3} \end{pmatrix} \quad e_1 = \begin{pmatrix} 0 & 0 \\ 0 & \frac{2}{3} \end{pmatrix}$$

$$U_1^{(2)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$U_2^{(2)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

So

$$U(\lambda) = I - \lambda \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \lambda^2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

4. Summary of Preceding Algorithms

We now summarize the algorithms described in this report. APL implementations for these algorithms can be found in [2] .

Algorithm 1. Given the polynomial matrix $P(\lambda)$ with linearly independent rows for every λ , this algorithm finds a polynomial matrix $Q(\lambda)$, such that

$$R(\lambda) = \begin{bmatrix} P(\lambda) \\ Q(\lambda) \end{bmatrix}$$

is unimodular.

1) Find constant non-singular matrices V_1 and V_2 such that

$$V_1 P(\lambda) V_2 = [I \quad 0] + \lambda P_1 + \ldots + \lambda^{\iota} P_{\iota}$$

2)

$$A_0 \triangleq A = \begin{bmatrix} 0 & I_m & 0 & & 0 \\ & & \ddots & & \vdots \\ & & & \ddots & 0 \\ & & & & I_m \\ \begin{bmatrix} -P_{\iota} \\ 0 \end{bmatrix} & & & & \begin{bmatrix} -P_1 \\ 0 \end{bmatrix} \end{bmatrix}, \quad B_0 \triangleq B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I_{m-n} \end{bmatrix}$$

$r_0 = \iota m$

$j = 0$

3) Find $\rho_j = \text{rank } (A_j)$

4) If $\rho_j = r_j$, set $a=j$ go to 7

5) Find $L_{j+1}(r_j \times \rho_j)$, $M_{j+1}(\rho_j \times r_j)$, of full column and row

rank, respectively, so that

$$A_j = L_{j+1}M_{j+1}$$

6) Set

$$A_{j+1} = M_{j+1}L_{j+1}$$

$$B_{j+1} = M_{j+1}B_j$$

$$r_{j+1} = P_j$$

$$j = j + 1$$

go to 3

7) If $r_a = 0$, $Q(\lambda) = [0_n \quad I_{n-m}]V_2^{-1}$, stop.

8) Find F_a such that $A + B F$ is nilpotent (see Appendix). Set

$$F = F_a M_a M_{a-1} \dots M_1$$

and partition

$$F = [-F_\ell \quad -F_{\ell-1} \dots -F_1]$$

$$Q(\lambda) = \left([0 \quad I_{n-m}] + \lambda F_1 + \dots + \lambda^\ell F_\ell\right) V_2^{-1}$$

<u>Algorithm 2</u>. Given a square nxn polynomial matrix

$$R(\lambda) = R_0 + \lambda R_1 + \dots + \lambda^\ell R_\ell$$

this algorithm checks if $R(\lambda)$ is unimodular and if so finds the inverse

$$U(\lambda) = R^{-1}(\lambda)$$

1) Set $V = R_0^{-1}$

and overwrite $R(\lambda) = VR(\lambda)$

2) Find

$$F_0 = D_0 = P_0 = \sum_{j=0}^{\ell} R_j^T R_j$$

$$e_0 = P_0 - R_0^T R_0$$

$$R_0 = I_n$$

Set

$$k = 1 \quad \text{and} \quad U^{(0)} = G^{(0)} = 0$$

3) $$P_k = \sum_{j=0}^{\iota-k} R_j^T R_{j+k} \qquad \text{if} \quad \iota \geq k$$

$$= 0 \qquad \text{if} \quad \iota < k$$

$$W = P_k - [P_{k-1} \ldots P_1] \; U^{(k-1)}$$

$$U_k^{(k)} = D_{k-1}^{-1} W$$

$$\begin{bmatrix} U_1^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ U_{k-1}^{(k)} \end{bmatrix} = \begin{bmatrix} U^{(k-1)} \\ \cdot \\ \cdot \\ \cdot \\ U_{k-1}^{(k-1)} \end{bmatrix} - \begin{bmatrix} G_{k-1}^{(k-1)} \\ \cdot \\ \cdot \\ \cdot \\ G_1^{(k-1)} \end{bmatrix} U_k^{(k)}$$

$$e_k = e_{k-1} - W^T U_k^{(k)}$$

ϵ_k = sum of the square roots of the diagonal elements of e .

4) If $\epsilon_k \leq \epsilon$ (for some small $\epsilon > 0$)

set $d = k$

$$U(\lambda) = \left[I_n - \lambda U^{(d)} - \ldots - \lambda^d U_d^{(d)} \right] V$$

Stop.

5) If $k \geq \iota n$

$R(\lambda)$ is not unimodular

Stop.

6) $$G_k^{(k)} = F_{k-1}^{-1} W^T$$

$$\begin{bmatrix} G_{k-1}^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ G_1^{(k)} \end{bmatrix} = G^{(k-1)} - U^{(k-1)} G_k^{(k)}$$

7)* $$D_k = P_o - [P_k \ldots P_1]\ G^{(k)}$$

and

$$F_k = P_o - [P_1^T \ldots P_k^T]\ U^{(k)}$$

8) Go to 3

---

* In step 7 $[P_k, \ldots, P_1]$ and $G^{(k)}$ are of order $nx(nk)$ and $nkxn$ respectively.

We will now show an alternative way for step 7 that saves some computation time. It is

$$\begin{aligned}
D_k &= P_o - [P_k \ldots P_1]\ G^{(k)} \\
&= P_o - [P_{k-1} \ldots P_1] \begin{bmatrix} G_{k-1}^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ G_1^{(k)} \end{bmatrix} - P_k G_k^{(k)} \\
&= P_o - [P_{k-1} \ldots P_1] \left\{ G^{(k-1)} - U^{(k-1)} G_k^{(k)} \right\} - P_k G_k^{(k)} \\
&= P_o - [P_{k-1} \ldots P_1] G^{(k-1)} + [P_{k-1} \ldots P_1] U^{(k-1)} G_k^{(k)} - P_k G_k^{(k)} \\
&= D_{k-1} - \left(P_k - [P_{k-1} \ldots P_1]\ U^{(k-1)} \right) G_{kk}^{(k)} \\
&= D_{k-1} - W G_k^{(k)}
\end{aligned}$$

where W has already been computed in step 3. Similarly we get

$$F_k = F_{k-1} - W^T U_k^{(k)}$$

5. Examples

The two algorithms in Section 3 were implemented in APL on the ITEL AS/6 system at Rice University. [2]

In the following examples a matrix $P(\lambda)$ with linearly independent rows is given. The matrices $Q(\lambda)$ and $U(\lambda)$ are found so that

$$U(\lambda) = \begin{pmatrix} P(\lambda) \\ Q(\lambda) \end{pmatrix}^{-1}$$

($\begin{pmatrix} P(\lambda) \\ Q(\lambda) \end{pmatrix}$ is unimodular) .

The results are checked as follows:

(a) The infinity norm of e_k , $\| e_k \|_\infty$ is tabulated for $k=1,2,\ldots$. The infinity norm of a matrix $A = [\alpha_{ij}]$ is defined by

$$\| A \|_\infty = \max_i \left| \sum_j |\alpha_{ij}| \right|$$

and

(b) The product $P(\lambda)U(\lambda)$ is formed. Recall that it should be $P(\lambda)U(\lambda) = [I \;\; 0]$.

Example 1

$$P(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \lambda + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \lambda^2$$

$$Q(\lambda) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & -1 & -1 \end{pmatrix} \lambda + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \lambda^2$$

$$U(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \lambda + \begin{pmatrix} 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \lambda^2$$

k	1	2
$\|e_k\|_\infty$	4.282	2.522E-14

$$P(\lambda)U(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

<u>Example 2</u>

$$\begin{aligned} P(\lambda) &= [\ .333 \quad .667] \\ &+ [\ .000 \quad .667]\ \lambda \\ &+ [1.667 \quad -1.000]\ \lambda^2 \\ &+ [\ .667 \quad -.333]\ \lambda^3 \\ &+ [\ .333 \quad .333]\ \lambda^4 \end{aligned}$$

$$\begin{aligned} Q(\lambda) &= [-.894 \quad .447] \\ &+ [\ .534 \quad -.720]\ \lambda \\ &+ [-.656 \quad .203]\ \lambda^2 \\ &+ [-.173 \quad .337]\ \lambda^3 \\ &+ [-.170 \quad -.170]\ \lambda^4 \end{aligned}$$

$$\begin{aligned} U(\lambda) &= \begin{pmatrix} .600 & -.894 \\ 1.200 & .447 \end{pmatrix} \\ &+ \begin{pmatrix} -.966 & -.894 \\ -.717 & .000 \end{pmatrix} \lambda \\ &+ \begin{pmatrix} .273 & 1.342 \\ .881 & 2.236 \end{pmatrix} \lambda^2 \\ &+ \begin{pmatrix} .452 & .447 \\ .232 & .894 \end{pmatrix} \lambda^3 \\ &+ \begin{pmatrix} -.228 & -.447 \\ .228 & .447 \end{pmatrix} \lambda^4 \end{aligned}$$

k	1	2	3	4
$\|e_k\|_\infty$	8.580	3.099	2.688	1.954E-14

$$PU = [1 \quad 0]$$

Example 3

$$P(\lambda) = [-.466 \quad .324 \quad -.052 \quad .042 \quad -.356] + [-.573 \quad .227 \quad .228 \quad .551 \quad -.147]\,\lambda$$

$$Q(\lambda) = \begin{bmatrix} .481 & .863 & .022 & -.018 & .150 \\ -.077 & .022 & .996 & .003 & -.024 \\ .062 & -.018 & .003 & .998 & .019 \\ -.529 & .150 & -.024 & .019 & .835 \end{bmatrix} + \begin{bmatrix} -.137 & .054 & .055 & .132 & -.035 \\ .362 & -.143 & -.144 & -.348 & .093 \\ .653 & -.259 & -.260 & -.628 & .168 \\ .282 & -.112 & -.112 & -.271 & .072 \end{bmatrix} \lambda$$

$$U(\lambda) = \begin{bmatrix} -1.027 & .481 & -.077 & .062 & -.529 \\ .715 & .863 & .022 & -.018 & .150 \\ -.114 & .022 & .996 & .003 & -.024 \\ .092 & -.018 & .003 & .998 & .019 \\ -.784 & .150 & -.024 & .019 & .835 \end{bmatrix} + \begin{bmatrix} .600 & -.072 & .190 & .343 & .148 \\ -.760 & .091 & -.240 & -.434 & -.187 \\ .650 & -.078 & .205 & .371 & .160 \\ .944 & -.113 & .299 & .539 & .233 \\ 1.038 & -.125 & .328 & .593 & .256 \end{bmatrix} \lambda$$

k	1
$\|\|e_k\|\|_\infty$	1.540E-15

$$PU = [1 \quad 0 \quad 0 \quad 0 \quad 0]$$

Example 4

$$P(\lambda) = \begin{pmatrix} .991 & -.689 & .110 & -.089 & .757 \\ 1.217 & -.482 & -.485 & -1.172 & .312 \end{pmatrix}$$

$$+ \begin{pmatrix} -.054 & -.891 & 1.252 & 1.201 & -.081 \\ -.463 & 1.325 & .312 & 1.166 & .221 \end{pmatrix} \lambda$$

$$+ \begin{pmatrix} -.504 & -.240 & -1.161 & -.935 & -.073 \\ 1.099 & -.415 & .226 & -.544 & -1.107 \end{pmatrix} \lambda^2$$

$$Q(\lambda) = \begin{pmatrix} .202 & .389 & .839 & -.316 & -.070 \\ .566 & .645 & -.282 & .423 & -.063 \\ -.311 & .436 & -.147 & -.229 & .799 \end{pmatrix}$$

$$+ \begin{pmatrix} -6.422 & 37.077 & -26.672 & -10.492 & -4.976 \\ .688 & 4.492 & -3.187 & -1.828 & -2.990 \\ -.607 & 1.869 & -.513 & .414 & .633 \end{pmatrix} \lambda$$

$$+ \begin{pmatrix} 32.946 & -5.622 & 23.541 & 2.457 & -23.992 \\ 5.054 & -1.212 & 2.752 & -.585 & -4.152 \\ 1.219 & -.149 & 1.016 & .253 & -.808 \end{pmatrix} \lambda^2$$

$$U(\lambda) = \begin{pmatrix} .317 & .187 & .202 & .566 & -.311 \\ -.402 & .074 & .389 & .645 & .436 \\ .343 & -.325 & .839 & -.282 & -.147 \\ .499 & -.609 & -.316 & .423 & -.229 \\ .549 & -.202 & -.070 & -.063 & .799 \end{pmatrix}$$

$$+ \begin{pmatrix} 9.234 & -4.899 & .971 & -6.522 & -3.652 \\ 17.403 & -9.257 & 2.276 & -12.298 & -9.199 \\ 26.584 & -13.737 & 4.663 & -18.210 & -16.423 \\ -9.243 & 4.233 & -1.735 & 7.035 & 7.518 \\ -2.955 & 2.067 & .503 & 1.410 & .864 \end{pmatrix} \lambda$$

$$+ \begin{pmatrix} -16.284 & 8.730 & -2.530 & 10.931 & 8.462 \\ -18.228 & 9.772 & -2.832 & 12.236 & 9.472 \\ 13.803 & -7.400 & 2.145 & -9.266 & -7.173 \\ -3.314 & 1.777 & -.515 & 2.225 & 1.722 \\ -4.885 & 2.619 & -.759 & 3.279 & 2.539 \end{pmatrix} \lambda^2$$

k	1	2
$\|e_k\|_\infty$	1716.187	1.992E-07

$$P(\lambda)U(\lambda) = \begin{pmatrix} 1.0E00 & 0.0E00 & 0.0E00 & 0.0E00 & 0.0E00 \\ 0.0E00 & 1.0E00 & 0.0E00 & 0.0E00 & 0.0E00 \end{pmatrix}$$

$$+ \begin{pmatrix} -5.6E-10 & 3.0E-10 & -9.5E-11 & 3.7E-10 & 3.0E-10 \\ -2.1E-09 & 1.1E-09 & -3.5E010 & 1.4E-09 & 1.1E-09 \end{pmatrix} \lambda$$

As expected, the error $\|e_k\|$ decreases and at the point $k=d$ drops to a "very small" value. Also, it seems there is a "discontinuity" in $\|e_k\|$ at the point $k=d$, in the sense that the number

$$K(d) = \begin{cases} \|e_d\| & , d=1 \\ \dfrac{\|e_d\|}{\|e_{d-1}\|} & , d=2 \\ \dfrac{\|e_d\|}{\|e_{d-1}\|} \Big/ \dfrac{\|e_{d-1}\|}{\|e_{d-2}\|} & , d>2 \end{cases}$$

is "very small".

The decision of what a "very small" number is depends upon the particular problem. Usually "a quantity is very small, and may be set to zero, if a perturbation of the same size can be tolerated in the original data". [6]

6. References

[1] Dickinson, B. W., Kailath, T. and Morf, M., "Canonical Matrix Fraction and State Space Descriptions for Deterministic and Stochastic Linear Systems", IEEE Trans. Autom. Control, Vol. AC-6, pp. 656-667, December 1974.

[2] Kontos, A.,"APL Programs for Polynomial Matrix Manipulations", Rice Univ. Tech. Report No. 7913, December 1979.

[3] Laub, A. J. and Moore, B. C., "Calculation of Transmission Zeros Using QZ Techniques", Automatica, Vol. 14, 1978.

[4] Levinson, L., Appendix B of N. Wiener, "Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications", pp. 129-148, John Wiley & Sons, New York, 1949.

[5] Luenberger, D. G., "Canonical Forms for Linear Multivariable Systems", IEEE Trans. Autom. Control, Vol. AC-12, pp. 290-293, 1967.

[6] Moler, C. B. and Stewart, G. W., "An Algorithm for Generalized Matrix Eigenvalue Problems", SIAM J. Numer. Anal., Vol. 10, pp. 241-256, April 1973.

[7] Moore, B. C. and Laub, A. J., "Computation of Supremal (A,B)-Invariant and Controllability Subspaces", IEEE Trans. Autom. Control, Vol. AC-23, pp. 783-792, October 1978.

[8] Rosenbrock, H. H., "State Space and Multivariable Theory", Nelson, London, 1970.

[9] Silverman, L. M. and Van Dooren, P., "A System Theoretic Approach for GCD Extraction", Proc. 1978 IEEE Conf. on Decision and Control, San Diego, California, January 10-12, 1979, pp. 525-528.

[10] Stewart, G. W., "Introduction to Matrix Computations", Academic Press, New York, 1973.

[11] Van Dooren, P., "The Generalized Eigenstructure Problem. Part I: Theory", USCEE Report 503, January 1979.

[12] Whittle, P., "On the Fitting of Multivariable Autoregressions, and the Approximate Canonical Factorization of a Spectral Density Matrix", Biometrica, Vol. 50, pp. 129-134, 1963.

[13] Wiggins, R. A. and Robinson, E. A., "Recursive Solution to the Multichannel Filtering Problem", J. Geophys. Res., Vol. 70, No. 8, pp. 1885-1891, April 1965.

[14] Wolovich, W. A., "Linear Multivariable Systems", Springer-Verlag, New York, 1974.

Appendix

In this appendix it will be shown how, given A(nxn) and B(nxm) , a matrix F(mxn) can be found such that A+BF is nilpotent. First let B_1 and L of column and row rank respectively, such that $B = B_1 L$. It is shown in [5] that a non-singular matrix T can be found* such that

$$\hat{A} \triangleq TAT^{-1} = A_1 + DG$$

and

$$\hat{B} \triangleq TB_1 = DK$$

where

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & . & . & 0 \\ & & . & . & . & 1 \\ 0 & . & . & . & . & 0 \end{bmatrix}$$

D is a matrix of the form

$$D = \begin{bmatrix} D_1 \\ D_2 \\ . \\ . \\ . \\ D_m \end{bmatrix}$$

with

$$D_j = \begin{bmatrix} 0 & . & . & . & 0 & . & . & . & 0 \\ & & & & . & & & & \\ & & & & . & & & & \\ & & & & . & & & & \\ 0 & . & . & . & 0 & . & . & . & 0 \\ 0 & & & 0 & 1 & 0 & . & . & 0 \end{bmatrix} , \quad j=1,2,\ldots,m$$

$$\uparrow$$
$$j^{th} \text{ column}$$

---

* An APL program that computes the matrix T is given in [2] .

i.e., all elements of D_j are zero except the last element of the j^{th} column which is 1 .

K is an mxm upper triangular matrix with 1's on the diagonal, and

G is some mxn matrix.

Define $\hat{F}$ through

$$K \hat{F} = - G$$

(Because of the special form of K , it is very easy to solve for $\hat{F}$). Since L has full row rank, there exists a matrix L^+ such that

$$L L^+ = I .$$

Define

$$F = L^+ \hat{F} T^{-1}$$

Then

$$\begin{aligned} A + BF &= A + B_1 L L^+ \hat{F} T \\ &\equiv T^{-1} A T + T^{-1} \hat{B} \hat{F} T \\ &= T^{-1} \left[\hat{A} + \hat{B}(-K^{-1} G) \right] T \\ &= T^{-1} \left[A_1 + DG - D K K^{-1} G \right] T \\ &= T^{-1} A_1 T , \end{aligned}$$

which is nilpotent.