April, 1976 Technical Memorandum ESL-TM-659

Recent Results on Observer-Based Compensation of Linear Systems

by

K. Tong and T.L. Johnson

This research has been conducted at the M.I.T. Electronic Systems Laboratory with support provided by the National Aeronautics and Space Administration NASA/Ames Grant NGL 22-009-124, and by the Esther and Harold E. Edgerton Professorship (second co-author).

Electronic Systems Laboratory
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Contents

I. Introduction

This technical memorandum presents a variety of recent results related to the design of observer-based compensators for linear constant systems, having dimension $s < n-m$, the number of states less the number of outputs. In the next section, the case $s \geq \max\,((n-m)r/m,\ \nu-1)$ is considered, and a method for decoupling the constraint equation of the observer is presented; this provides an extension and correction of Rothschild's results. This condition on s is not necessary and sufficient, however, and an appropriate modification of Miller's results is presented in the Addendum. The condition is directly-related to some requirements stated by Seragi, for the frequency-domain design of pole-placement compensators, and this relation is explored in Section IV.

One approach to the solution of necessary conditions for quadratic-cost optimal compensators involves symbolic solution of certain Lyapunov equations. The applicability of Kronecher-product representations for the discrete and continuous-time version of such equations to various compensation problems is shown in Sections V and VI. Some worked examples are given in Section VII, using the symbolic-inversion capabilities of the MACSYMA.

We note that these results are of a preliminary nature, and may be subject to further refinement. In particular, the questions of existence and uniqueness remain to be resolved.

Section II.

The below-minimal order asymptotic reconstructor of r linear functions of state $\underline{x}$. (Rothschild's result extended and corrected).

Plant: $\dot{x} = Ax + Bu,\ x(0) = x_0$ (1)

$y = Cx$, rank $(C) = m < n$ (2)

$u = Fx$, rank $(F) = r < n$ (3)

Asymptotic Compensator:

$\dot{z} = Dz + Ey + Gu$ (4) [s-order Luenberger observer, s<n-m

$z \to Tx$ asymptotically (Miller's notation]

In fact,

$z = Tx + \exp(Dt)(z(0)-Tx(0))$ (5)

For this equation to be true, Luenberger worked out the following constraint, hereafter referred to as the Luenberger condition, on D and E.

$TA - DT = EC$ (6) $G = TB$ (6a)

(It has a unique solution for T when A and D have no eigenvalues in common. When they do, solution also exists, although not unique.) An estimate of u is constructed using y and z as follows

$$\hat{u} = Hy + Mz \qquad (7)$$

and we require $\hat{u} \to u$ asymptotically. Putting (2) and (5) into (7), we obtain

$$\hat{u} = HCx + MTx + M\exp(DT)(z(0)-Tx(0))$$

$$= Fx + M\exp(DT)(z(0)-Tx(0)) \qquad (8)$$

$$\text{if } HC + MT = F \qquad (9)$$

When (9) is satisfied, $\hat{u}$ does approach u asymptotically and we refer to (9) as the asymptotic condition.

Below-min. order vs Min. order compensator

As long as we are only concerned with $\hat{u} \rightarrow u$, our job is to fix F and determine D (has to be strickly stable), E, T, H and M in (4) and (7) so that (6) and (9) are satisfied. G in (4) is fixed by (6a). One sequence of steps (Design #1) to accomplish this is to fix D (hence pole placement of the open-loop compensator) and E so that (D,E) is a controllable pair and solve (6) for T. A solution always exists because (6) has as many linear equations as there are elements in T. When T is non-unique, pick one sol. arbitrarily, put it in (9) and solve for H and M. Equation (9) has rn linear equations in rm elements in H and rs elements in M. A necessary condition for Eq. (9) to have a solution is that there are at least as many unknowns as there are equations.

$$rm + rs \geq rn$$

$$\text{or} \quad s \geq n-m \tag{10}$$

Alas, we only design a min. order asymptotic compensator of order at least n-m. The following sequence of steps (Design #2), however, will permit us to design a compensator with

$$s \geq \max\left(\frac{n-m}{m} r,\ \nu-1\right) \triangleq \gamma \tag{11}$$

where ν is the observability index of (A,C) and r is the dimension of the control vector u. Since the observability index is bounded by

$$\frac{n}{m} \leq \nu \leq n-m+1 \qquad (12)$$

We will show in the following that γ has a max. at n-m and hence a below-min. order compensator is possible. γ can take on either arguments in (11) depending on which is bigger.

- when $\gamma = \nu-1$ (i.e. $\nu-1 < \frac{n-m}{m} r$)

$\therefore \gamma = \nu-1 \leq n-m$, by (12)

- when $\gamma = \frac{n-m}{m} r$ (i.e. $\frac{n-m}{m} r > \nu-1$
  - when $m \geq r$, $\leq n-m$
  - when $m < r$, impossible because if this is true, $\gamma \geq n-m$ which contradicts the assumption that

    $\underbrace{\gamma = \frac{n-m}{m} r > \nu-1}_{\text{assumption}} \geq \frac{n}{m} - 1 = \frac{n-m}{m}$ by (12)

    assumption

    contradiction when $m < r$

□

The below-min. order compensator design sequence (Design #2) is as follows. Fix D and M so that (D,M) is an observable pair. Solve (6) and (9) simultaneously for H, E. and T. We indeed have a price to pay to go below-min. order because whereas design #1 solves (6) and (9) successively, we now have to do it simultaneously. Luckily, Rothschild and Jameson have devised a method (to be described later) that decouples the requirement of simultaneity. So far, we have not shown why fixing D and M as opposed to fixing D and E will allow us to go below min. order. The answer lies in the number of elements in D and E. More specifically, (6) and (9) have sn and rn linear equations and H, E and T have rn, sm and sn elements respectively. A <u>necessary</u> condition for sol. to H, E. and T to exist is there are at least as many unknowns as there are linear equations.

i.e. $rm + sm + sn \geq sn + rn$

or $$s \geq \frac{n-m}{m} r \tag{13}$$

This is the first argument in (11). The other argument comes from the sufficient condition for existence of solution, is tied up with the rank of the coefficient matrix of the sn + rn equations, is related to the decoupling of simultaneity mentioned before the observability of (D,M) and will be discussed in due course. Now, compare the necessary conditions (13) and (10). When $r<m$, it is obvious that design #2 may go below min. order. When $r>m$, however, the sufficient condition reflected in the 2nd argument in (11) places n-m as an upper bound to γ as we have seen and again implies design #2 may go below min. order. This is not quite fair because we have not mentioned the sufficient condition for sol. in design #1. Just as the sufficient condition for design #2 is related to the observability of (D,M), the sufficient condition for design #1 has to do with the controllability of (D,E). But this is a red herring because if $s \geq n-m$ is necessary for design #1, the sufficient condition can only push min. s up and not down, otherwise the necessary condition is violated (and we do not have a solution).

We will now show how the design to min. a quadratic performance index using a min. order Luenberger observer as solved by Miller, etc. really uses the philosophy of design #1. We emphasize "philosophy" because design by min.

cost and by pole placement are related but not, naturally, the same. Their relation is discussed in the next section.

It is not difficult to show that, as Newmann 1970 has done, D = TAM, E = TAH are necessary and sufficient for the Luenberger condition (6) to be satisfied when s = n-m. (When s > n-m, they are only sufficient; when s < n-m, they are neither necessary nor sufficient.) Hence D and E are fixed in relation to M and H. The cost minimization proceeds to consider M, H and T free except that (9) (in a slightly modified form) has to be satisfied. This is exactly the philosophy of design #1. In fact, $[H\ M] = \begin{bmatrix} C \\ T \end{bmatrix}^{-1}$. We mentioned previously that design #1 requires (D, E) to be a controllable pair. That (TAM, TAH) is controllable requires (A, C) to be observable. Observability of (A, C) is a necessary condition for a Luenberger observer to exist. Since B does not affect observability, set it to zero temporarily in the following demonstration:

$$\left.\begin{aligned} \dot{x} &= Ax \\ y &= Cx \quad , \quad (A, C) \text{ observable} \end{aligned}\right\} \tag{14}$$

For $s = n\text{-}m$, $\begin{bmatrix} H' \\ M' \end{bmatrix} = (C'\ T')^{-1}$ is non-sing. Apply the non-singular transformation

$$\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} = \begin{pmatrix} H' \\ M' \end{pmatrix} x \tag{15}$$

$$\begin{pmatrix} H' \\ M' \end{pmatrix}^{-1} = (C'\ T') \tag{16}$$

to (14) and we get

$$\left.\begin{aligned} \frac{d}{dt}\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} &= \begin{pmatrix} H'AC' & H'AT' \\ M'AC' & M'AT' \end{pmatrix}\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} \\ y &= C(C'T')\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} \end{aligned}\right\} \qquad (17)$$

From Lemma 2 of Luenberger 1971 (IEEE Trans., Dec.), (A,C) observable $\Rightarrow$ (A_{22}, A_{12}) observable. Since observability is not altered, (M'AT', H'AT') is observable $\Leftrightarrow$ (TAM, TAH) is controllable.

That D = TAM is strictly stable requires $\sum_o = E[(x_o-E(x_o))(x_o-E(x_o))'] > 0$ and (A,C) observable. This is proved in Lemma 5 in Miller, 1972.

A corollary is as follows.* In designing a below min. order observer that min. a performance index (extension of Miller's result to $s < n-m$) by assuming $D = TAF^+M$, $E = TAF^+H$, F^+ = left inverse of F, i.e. $F^+F = I$, so as to satisfy (6) identically and proceed to opt. wrt M, H and T under constraint (9) is bound to end in failure because it uses the design #1 philosophy which requires $s \geq n-m$, the implication is we have more freedom by using a lower order compensation, an apparent contradiction.

Pole-placement and minimum cost designs

The designs #1 and #2 in the last section are pole-placement designs. In this section, we will show their relation to min. cost design. The bridge between the two is in the cost separation lemma which is proved in Newmann 1969

---

*See addendum following; a modification of this procedure will in fact work.

which is valid for any $s \geq 1$. The cost separation lemma assumes the separation theorem to hold so that $\hat{u} = F\hat{x}$($\hat{u}$=$F\hat{x}$ for $s \geq n-m$) is used to replace $u = Fx$, F is the optimal gain when x is completely accessible. The lemma is as follows.

<u>The cost separation lemma</u>

<u>$s \geq n-m$ version</u>

For a plant

$$\dot{x} = Ax + Bu$$

$$y = Cx, \quad \text{rank } (C) = m < n$$

and an observer

$$\dot{z} = Dz + Ey + Gu$$

$$TA - DT = EC, \quad G = TB, \tag{18}$$

so that u is estimated by

$$\left.\begin{aligned} \hat{u} &= F\hat{x} \\ \hat{x} &= Hy + Mz, \end{aligned}\right\} \tag{19}$$

the optimal D, E, H, M and T that minimize a quad. perf. index

$$J = \int_0^\infty (x'Qx + u'Ru)\,dt$$

is obtained by solving the optimization part

$$\min_{M,T} \Delta J = \min_{M,T} (z(0) - Tx(0))' P_{22} (z(0) - Tx(0)) \tag{20}$$

where P_{22} satisfies the Lyaponov equation

$$D'P_{22} + P_{22}D = -M'F'RFM \qquad (21)$$

(The minimization is wrt M, T only because with D = TAM, only M and T are involved in the cost and the constraint) and the constraint part (for H)

$$HC + MT = I \qquad (22)$$

The optimum E will be given by TAH.

<u>s < n-m version</u>

(19) is replaced by

$$\hat{u} = \widehat{Fx} = Hy + Mz \qquad (23)$$

(21) is replaced by

$$D'P_{22} + P_{22}D = -M'RM \qquad (24)$$

(22) is replaced by

$$HC + MT = F \qquad (25)$$

D and E are no longer given by TAM and TAH respectively, so (20) is replaced by

$$\min_{D,M,T} \Delta J = \min_{D,M,T} (z(0) - Tx(0))' P_{22}(z(0) - Tx(0)) \qquad (26)$$

It is not hard to see that the s < n-m version can be extended to the s $\geq$ n-m version but not vice-versa. (Put F = I in (25), (21) and (19) will equate (19), (21), (22) to (23), (24) and (25) respectively. Furthermore set D, E to TAM, TAH). From here on, we mean the s < n-m version when we refer to the cost separation theorem.

* * *

The relation of the min. cost design to the pole-placement design is simply the addition of the optimization part (26) subjected to (24); and while D is free (by definition of pole-placement), so is M so long as (D,M) is observable, in the pole-placement design, they are to be optimized in the min. cost design--or one can say (D,M) is fixed by design and optimization respectively. In (26), we are very general and include T in the minimization. Actually, the optimal T can be solved from (18) and (25) together with the optimal H and E once D and M are fixed at their optimal values. In this light, we see (24) and (26) really use the design #2 philosophy and hence below-min. order compensator is possible.

<u>Rothschild-Jameson's method to decouple (18) and (25) into 2 matrix equations, one involving T, the other involving only H and E</u>

That T appears in both (18) and (25) provides a coupling between the two which requires their simultaneous solution. A decoupling that provides for sequential solution is not only more attractive computationally, but also provides the condition for existence of solution for H and E, which formerly appear in 2 different matrix equations.

For ease of exposition, we will first consider the single-input case,

$r = 1$. The multiple-input case is a straightforward extension.

$\underline{r = 1}$: Since H, M & F are row vectors now, denote them by $\underline{h}^T$, $\underline{m}^T$ & $\underline{f}^T$. (18) and (25) will become

$$TA - DT = EC \tag{26}$$

$$\underline{h}^T C + \underline{m}^T T = \underline{f}^T \tag{27}$$

The characteristic polynomial of the compensator is

$$\det(\lambda I - D) = \lambda^s + d_1 \lambda^{s-1} + \ldots + d_s = 0 \tag{28}$$

Define

$$\left.\begin{aligned}
C_0 &\triangleq 0 \\
C_1 &\triangleq EC = TA - DT \text{ by } (26) \\
C_2 &\triangleq ECA + DC_1 = TA^2 - D^2T \text{ by } C_1 \ \& \ (26) \\
C_3 &\triangleq ECA^2 + DC_2 = TA^3 - D^3T, \text{ by } C_2 \ \& \ (26) \\
&\vdots \\
C_i &\triangleq ECA^{i-1} + DC_{i-1} = TA^i - D^iT \quad , \quad s \geq i \geq 1
\end{aligned}\right\} \tag{29}$$

Multiply each C_i by d_{s-i} and sum we obtain

$$\sum_{i=0}^{s} Td_i A^{s-1} = \sum_{i=0}^{s} d_i C_{s-i} + \sum_{i=0}^{s} d_i D^{s-i} T \tag{30}$$

where $d_0 = 1$. From the Cayley-Hamilton theorem, D satisfies its own characteristic equation,

$$\sum_{i=0}^{s} d_i D^{s-i} = 0, \tag{31}$$

so (30) reduces to

$$T(A^2 + d_1 A^{s-1} + \ldots + d_s I) = C_s + d_1 C_{s-1} + \ldots + d_{s-1} C_1 \tag{32}$$

Expanding the right-hand side according to definitions of C_i's, we have

$$T(A^2 + d_1 A^{s-1} + \ldots + d_s I) = ECA^{s-1} + (D + d_1 I)ECA^{s-2} + (D^2 + d_1 D + d_2 I)ECA^{s-3}$$

$$+ \ldots + (D^{s-1} + d_1 D^{s-2} + \ldots + d_{s-1} I)EC \tag{33}$$

Multiplying (33) on the left by $\underline{m}^T$ and using (27), we have

$$\underline{h}^T C(A^s + d_1 A^{s-1} + \ldots + d_s I) + \underline{s}_1^T ECA^{s-1} + \underline{s}_2^T ECA^{s-2} + \ldots + \underline{s}_s^T EC$$

$$= \underline{f}^T (A^s + d_1 A^{s-1} + \ldots + d_s I) \tag{34}$$

where $\underline{s}_i$ is just a short hand notation that represents the ith column of the S matrix defined by

$$S \triangleq [\underline{m} \vdots (D^T + d_1 I)\underline{m} \vdots \ldots\ldots \vdots (D^{T^{(s-1)}} + d_1 D^{T^{(s-2)}} + \ldots + d_{s-1} I)\underline{m}] \tag{35}$$

Now the term that postmiltiplies T in (33) is the characteristic equation of D in A, i.e. with D replaced by A in (31). Under the assumption that A and D do not have any common eigenvalues, this term will never be zero. Furthermore, Gantmacher (1960) shows that under the above assumption, (26) has a unique solution for T. This implies (33) has a unique solution and hence $(A^s + d_1A^{s-1} + \ldots + d_sI)$ has an inverse. Therefore, when A and D do not have any common eigenvalues, we can write (33) as

$$T = [ECA^{s-1} + (D + d_1I)ECA^{s-2} + \ldots + (D^{s-1} + d_1D^{s-2} + \ldots + d_{s-1}I)EC] \times [A^s + d_1A^{s-1} + \ldots + d_sI]^{-1} \qquad (36)$$

(34) and (36) are the two decoupled matrix equations that can be proved (see below) to be equivalent to the two coupled equations (26) and (27). So instead of solving (26) and (27) simultaneously in T, $\underline{h}^T$ and E, we only have to solve (34) in $\underline{h}^T$ and E and using the E thus obtained to get T directly from (36).

<u>Proof</u> That (26) and (27) $\Rightarrow$ (34) and (36) is true is by construction. To show that (34) & (36) $\Rightarrow$ (27), we note that the value of $(\underline{f}^T-\underline{h}^TC)$ obtained from (34) must equal $\underline{m}^T$ multiplied by T. Indeed this is true when T is given by (36), after we consider the definition of $\underline{s}_i$ and equation (35).

To show that (36) $\Rightarrow$ (26), we first show (33) $\Rightarrow$ (26). Postmultiply (33) by A and subtract from the product (33) postmultiplied by D.

$$T(A^s + d_1A^{s-1} + d_2A^{s-2} + \ldots + d_sI)A - DT(A^2 + d_1A^{s-1} + d_sA^{s-2} + \ldots + d_sI)$$

$$= ECA^s + (D + d_1I)ECA^{s-1} + (D^2 + d_1D + d_2I)\ ECA^{s-2}$$

$$+ (D^3 + d_1D^2 + d_2D + d_3I)ECA^{s-3} + \ldots + (D^{s-1} + d_1D^{s-2} + \ldots + d_{s-1}I)ECA$$

$$- DECA^{s-1} - (D^2 + d_1D)ECA^{s-2} - (D^3 + d_1D^2 + d_2D)ECA^{s-3}$$

$$- \ldots - (D^{s-1} + d_1D^{s-2} + \ldots + d_{s-2}D)\ ECA$$

$$-(D + d_1D^{s-1} + \ldots + d_{s-1}D)\ EC \tag{37}$$

After noting the cancellations of the up-arrows with the down-arrows, the RHS of (37) is

$$EC(A^s + d_1A^{s-1} + d_2A^{s-2} + d_3A^{s-3} + \ldots + d_{s-1}A)$$

$$-(D + d_1D^{s-1} + \ldots + d_{s-1}D)EC \tag{38}$$

Add and subtract $EC\,d_s$ to each of the two terms above, to get

$$EC(A^s + d_1A^{s-1} + d_sA^{s-2} + d_3A^{s-3} + \ldots + d_{s-1}A + d_sI) - \cancel{ECd_s}$$

$$- (D + d_1D^{s-1} + \ldots + d_{s-1}D + d_sI)EC + \cancel{ECd_s} \tag{39}$$

Noting (31), the 2nd term of (39) is zero. Equating the LHS of (37) to the RHS of (39) and noting that $A^sA = AA^s$, etc.,

$$TA(A^s + d_1A^{s-1} + d_sA^{s-2} + \ldots + d_sI)A - DT(A^s + d_1A^{s-1} + d_2A^{s-2}$$

$$+ \ldots + d_sI) = EC(A^s + d_1A^{s-1} + d_2A^{s-2} + \ldots + d_sI) \qquad (40)$$

The assumption that $(A^2 + d_1A^{s-1} + d_2A^{s-2} + \ldots + d_sI)^{-1}$ exists proves (33)$\Rightarrow$(26). Since the same assumption proves (33)$\Leftrightarrow$(35), (35)$\Rightarrow$ (26). $\square$

Now let us examine the condition under which (34) has a solution to $\underline{h}^T$ and E. If we denote the s rows of E by $\underline{e}_1^T, \underline{e}_2^T, \ldots, \underline{e}_s^T$, the transpose of (34) can be written as

$$(A^{T^s} + d_1A^{T^{s-1}} + \ldots + d_sI)C^T\underline{h} + (s_{11}A^{T^{s-1}} + s_{12}A^{T^{s-2}} + \ldots + s_1\ I)C^T\underline{e}_1$$

$$+ (s_{21}A^{T^{s-1}} + s_{22}A^{T^{s-2}} + \ldots + s_{2s}I)C^T\underline{e}_2 + \ldots$$

$$+ (s_{s1}A^{T^{s-1}} + s_{s2}A^{T^{s-2}} + \ldots + s_{ss}I)C^T\underline{e}_s = (A^{T^s} + d_1A^{T^{s-1}} + \ldots + d_sI) \qquad (41)$$

To write (41) as a set of linear simultaneous equations $A_1\underline{x} = \underline{b}_1$ where $\underline{x}$ is the unknown vector $(\underline{h}^T|\underline{e}_1^T|\underline{e}_2^T|\ldots|\underline{e}_s^T)^T$, we have to define a matrix sm x sm matrix P:

$$P \triangleq S^T \times I_m = \begin{bmatrix} s_{11}I_m & s_{21}I_m & \cdots & s_{s1}I_m \\ s_{12}I_m & s_{22}I_m & \cdots & s_{s2}I_m \\ \vdots & \vdots & & \\ s_{1s}I_m & s_{2s}I_m & \cdots & s_{ss}I_m \end{bmatrix} \tag{42}$$

By multiplying out (43), it is not difficult to see that (41) can be written as

$$\underbrace{n \updownarrow \underbrace{\overset{\longleftarrow (s+1)m \longrightarrow}{[A^{T^s}C^T \vdots A^{T^{s-1}}C^T \vdots \cdots \vdots C^T]}}_{\triangleq \underline{A}} \overbrace{\begin{bmatrix} I_m & O_{m \times sm} \\ d_1 I_m & \\ d_2 I_m & \\ \vdots & P \\ d_s I_m & \end{bmatrix}}^{\triangleq B \atop \longleftarrow (s+1)m \longrightarrow}}_{\triangleq A_1} \underbrace{\begin{pmatrix} \underline{h} \\ \underline{e}_1 \\ \underline{e}_2 \\ \vdots \\ \underline{e}_s \end{pmatrix}}_{\underline{x}} = \underbrace{(A^{T^s} + d_1 A^{T^{s-1}} + \cdots + d_s I)\underline{f}}_{\underline{b}_1} \tag{43}$$

A necessary and sufficient condition for (43) to have a solution in $\underline{x}$ is

$$\text{rank}\ (A_1) = \text{rank}\ (A_1 \vdots \underline{b}_1) \tag{44}$$

or, stated in terms of vector space language, the columns of A_1 span $\underline{b}_1$. Since $\underline{b}_1$ is complicated function of $\underline{f}$, our following analysis considers $\underline{b}_1$ as a general vector, i.e., its elements can take any finite values. If $\underline{b}_1$ is a general vector, a necessary and sufficient condition for (43) to have a solution* is

*However, this solution may not be unique. See addendum. In general, some parameters of $\underline{h}$, $\underline{E}$ will be undetermined, and hence the following conditions are somewhat too conservative.

$$\text{rank } (A_1) = n \tag{44a}$$

(See C. T. Chen, Introduction to Linear Syst. Theory, p. 31, Thm. 2-4)

To determine the rank of A_1, we first have to determine the rank of S defined in (35). By definition, the observability matrix of $(D, \underline{m}^T)$ is

$$\mathcal{O}_{D,\underline{m}^T} \triangleq [\underline{m} \vdots D^T\underline{m} \vdots D^{T^2}\underline{m} \vdots \cdots \vdots D^{T^{s-1}}\underline{m}] \tag{45}$$

which has rank s iff $(D,\underline{m}^T)$ is observable. By elementary column operations on $\mathcal{O}_{D,\underline{m}^T}$, we can obtain S. E.g., the 2nd block of S is obtained by the sum of the 1st block of $\mathcal{O}_{D,\underline{m}^T}$ and d_1 times its 2nd block.

$$\therefore \text{ rank } (S) = \text{rank } (\mathcal{O}_{D,\underline{m}^T}) \triangleq \mu \tag{46}$$

Further more, $\mu = s$ iff $(D,\underline{m}^T)$ is observable. From the definition in (42), we can write

$$\text{rank } (P) = m \cdot \text{rank}(S) = m\mu \tag{47}$$

The observability index of (A,C) is defined as the least integer ν such that

$$\text{rank } [A^{T^{\nu-1}} C^T \vdots A^{T^{\nu-2}} C^T \vdots \cdots \vdots C^T] = n \tag{48}$$

We see that $s \geq \nu-1$ is necessary and sufficient that $\underline{A}$ has rank n.

Assume $s \geq \nu-1$ and $(D,\underline{m}^T)$ is observable. From (47) and the definition of $\underline{B}$ in (43), rank $(\underline{B}) = (s+1)m$. Hence $\underline{B}^{-1}$ exists. Since $A_1 = \underline{A}\,\underline{B}$, rank (A_1) = rank $(\underline{A}) = n$, by (48) and $s \geq \nu-1$. Together with (44a), we have just proved the sufficiency part of Theorem 1.

Theorem 1 An asymptotic estimate for a single input plant exists if $(D,\underline{m}^T)$ is observable and $s \geq \nu-1$. Given $(D,\underline{m}^T)$ is observable and $s \geq \nu-1$, the corresponding necessary condition is $s \geq \frac{n-m}{m}$. Since $\frac{n-m}{m}$ is always a lower bound to $\nu-1$, the sufficient conditions are also necessary.

Proof necessary condition. For $A_1 = \underline{A}\,\underline{B}$, it is necessary that rank $(A_1) \leq$ min (rank $(\underline{A})$, rank $(\underline{B})$) for (43) to have a solution. Given $(D,\underline{m}^T)$ is observable, rank $(\underline{B}) = (s+1)m$. Given $s \geq \nu-1$, rank $(\underline{A}) = n$. rank $(A_1) \leq \min(n,(s+1)m)$. But rank $(A_1) = n$, since the sufficient conditions are satisfied. $\therefore$ $n \leq \min(n,(s+1)m)$. Consider two cases: 1. $n \leq (s+1)m$ 2. $n > (s+1)m$.

1. $\min(n,(s+1)m) = n$ necessary condition is satisfied.
2. $\min(n,(s+1)m) = (s+1)m$. For the necessary condition to be true, $n \leq (s+1)m$ which contradicts the assumption of 2. $\therefore$ 2. can never satisfy the necessary condition.

Rearrange the assumption of 1. $s \geq \frac{n-m}{m}$. This will ensure that the necessary condition is satisfied. □

For a given s, an observable pair $(D,\underline{m}^T)$ can always be found. We can even say that for a given s and D (fixed by pole-placement), an $\underline{m}^T$ can always be found such that $(D,\underline{m}^T)$ is observable. (Convert D to observable canonical form, then set $\underline{m}^T$ to $(0,\ldots, 0, 1)$.)

When $(D,\underline{m}^T)$ is unobservable, rank $(\mathcal{O}_{D,\underline{m}^T}) < s$, rank $(S) < s$, rank $(P) < sm$, rank $(\underline{B}) < m+sm$, rank $(A_1) \leq \min(n, m+sm)$ and a solution to (43) is not guaranteed even when $s \geq \frac{n-m}{m}$ because if there are less than n independent equations in (43), the column space in A_1 may not span the vector $\underline{b}_1$. Rothschild has a necessary condition that guarantees a solution to (43) when $(D,\underline{m}^T)$ is unobservable. It is $s \geq \max(\nu-1, \frac{n-m}{m}, \frac{n}{m})$. This appears incorrect due to an error of reasoning in his two-line derivation.

<u>Extension to multiple-input case, r > 1</u>

There is no conceptual difficulty in extending Theorem 1 to apply to a multiple-input plant. We only have to deal with a bigger A_1 matrix, now called A_2 (also a bigger $\underline{x}$ and $\underline{b}_1$).

We start by decoupling the two matrix equations

$$TA - DT = EC \tag{54}$$

$$HC + MT = F \tag{55}$$

We will still keep (28), (29), (30), (31), (32), (33), (36), (37), (38), (39), and (40) which are independent of r. Multiplying (33) on the left by M and using (55), we have

$$HC(A^s + d_1A^{s-1} + \ldots + d_sI) + S^1ECA^{s-1} + S^2ECA^{s-2} + \ldots + S^sEC$$

$$= F(A^s + d_1A^{s-1} + \ldots + d_sI) \tag{56}$$

where S^i are compoenent blocks of S_r^T:

$$S_r \triangleq (s^{1^T} \mid s^{2^T} \mid \ldots \mid s^{s^T})$$

$$= (M^T \mid (D^T + d_1 I)M^T \mid (D^{T^2} + d_1 D^T + d_2 I)M^T \mid \ldots \mid (D^{T^{s-1}} + d_1 D^{T^{s-2}} + \ldots + d_{s-1} I)M^T) \tag{57}$$

To examine the condition under which (56) has a solution to H and E, denote the r rows of H, the s rows of E and the r rows of F by $\underline{h}_1^T$, $\underline{h}_2^T, \ldots, \underline{h}_r^T$; $\underline{e}_1^T$, $\underline{e}_2^T, \ldots, \underline{e}_s^T$, and $\underline{f}_1^T$, $\underline{f}_2^T, \ldots, \underline{f}_r^T$. Furthermore, as a short-hand, define

$$\Psi = A^s + d_1 A^{s-1} + \ldots + d_s I \tag{58}$$

Using a property of the Knonecker product that AXB = Y can be written as $(A \otimes B^T)\underset{\sim}{x} = \underset{\sim}{y}$ where $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are the <u>rows</u> of X and Y stringed out in a <u>column</u> vector, we can write (56) as

$$(I_r \otimes (C\Psi)^T)\begin{pmatrix} \underline{h}_1 \\ \vdots \\ \underline{h}_r \end{pmatrix} + [s^1 \otimes (CA^{s-1})^T + s^2 \otimes (CA^{s-2})^T + \ldots$$

$$+ s^s \otimes C^T]\begin{pmatrix} \underline{e}_1 \\ \vdots \\ \underline{e}_s \end{pmatrix} = (I_r \otimes \Psi^T)\begin{pmatrix} \underline{f}_1 \\ \underline{f}_2 \\ \vdots \\ \underline{f}_r \end{pmatrix} \tag{59}$$

which can be written in the form $A_2\,\underline{x} = \underline{b}$ as follows: Using the definition of a Knonecker product, we can write (59) as

$$
nr \updownarrow \underbrace{\begin{pmatrix} \Psi^T C^T & & 0 \\ & \ddots & \\ 0 & & \Psi^T C^T \end{pmatrix}}_{mr} \begin{pmatrix} \underline{h}_1 \\ \underline{h}_2 \\ \vdots \\ \underline{h}_r \end{pmatrix} + nr \updownarrow \underbrace{\begin{pmatrix} \theta_{11} & \theta_{12} & \cdots & \theta_{1s} \\ \theta_{21} & \theta_{22} & \cdots & \theta_{2s} \\ \vdots & \vdots & & \vdots \\ \theta_{r1} & \theta_{r2} & & \theta_{rs} \end{pmatrix}}_{ms} \begin{pmatrix} \underline{e}_1 \\ \underline{e}_2 \\ \vdots \\ \underline{e}_s \end{pmatrix}
$$

$$
= nr \updownarrow \underbrace{\begin{pmatrix} \Psi^T & & 0 \\ & \ddots & \\ 0 & & \Psi^T \end{pmatrix}}_{nr} \begin{pmatrix} \underline{f}_1 \\ \underline{f}_2 \\ \vdots \\ \underline{f} \end{pmatrix} \tag{60}
$$

where θ_{ij} is an n x m matrix given by

$$
\theta_{ij} = (S^1_{ij} A^{T^{s-1}} + S^2_{ij} A^{T^{s-2}} + \ldots + S^s_{ij} I) C^T
$$

and S^k_{ij} = ij-th element of S^k.

We can factor out $(A^{T^s} C^T \vdots A^{T^{s-1}} C^T \vdots \ldots \vdots C^T)$ from θ_{ij}:

$$
n \updownarrow \underbrace{[A^{T^S} C^T \vdots A^{T^{s-1}} C^T \vdots \ldots \vdots C^T]}_{(s+1)m} \begin{pmatrix} 0_m \\ S^1_{ij} I_m \\ S^2_{ij} I_m \\ \vdots \\ S^s_{ij} I_m \end{pmatrix} = \theta_{ij} \tag{61}
$$

Noting that $\Psi^T C^T$ can also be factored, we get

$$[A^{T^S} C^T \vdots A^{T^{s-1}} C^T \vdots \ldots \vdots C^T] \begin{pmatrix} I_m \\ d_1 I_m \\ d_2 I_m \\ \vdots \\ d_s I_m \end{pmatrix} = \Psi C^T \tag{62}$$

Define

$$n \updownarrow \overset{\leftrightarrow}{\underset{m}{\Gamma}} \triangleq [A^{T^S} C^T \vdots A^{T^{s-1}} C^T \vdots \ldots \vdots C^T] \tag{63}$$

$$(s+1)m \updownarrow \overset{\leftrightarrow}{\Delta} \triangleq \begin{pmatrix} I_m \\ d_1 I_m \\ d_2 I_m \\ \vdots \\ d_s I_m \end{pmatrix} \tag{64}$$

and

$$(s+1)m \updownarrow \overset{\overset{m}{\longleftrightarrow}}{\Sigma_{ij}} \triangleq \begin{pmatrix} 0_m \\ S_{ij}^1 \; I_m \\ S_{ij}^2 \; I_m \\ \vdots \\ S_{ij}^s \; I_m \end{pmatrix} \qquad \begin{array}{l} i = 1,2,\ldots, r \\ j = 1,2,\ldots, s \end{array} \tag{65}$$

(60) can be written as

$$
nr\;\underbrace{\underbrace{\overset{\longleftarrow\, r(s+1)m\,\longrightarrow}{\begin{pmatrix}\Gamma & & & 0\\ & \Gamma & & \\ & & \ddots & \\ 0 & & & \Gamma\end{pmatrix}}}_{\triangleq A_3}\;\underbrace{\left[\begin{array}{cccc|cccc} \multicolumn{4}{c|}{\longleftarrow rm \longrightarrow} & \multicolumn{4}{c}{\longleftarrow sm \longrightarrow}\\ \Delta & & & & \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{1s}\\ & \Delta & & & \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{2s}\\ & & \ddots & & \vdots & \vdots & \vdots & \vdots\\ & & & \Delta & \Sigma_{r1} & \Sigma_{r2} & \Sigma_{r3} & \Sigma_{rs}\end{array}\right]}_{\triangleq A_4}}_{\triangleq A_2}\cdot\begin{bmatrix}\underline{h}_1\\ \underline{h}_2\\ \vdots\\ \underline{h}_r\\ \cdots\\ \underline{e}_1\\ \underline{e}_2\\ \vdots\\ \underline{e}_s\end{bmatrix}
$$

$$
= \begin{pmatrix}\Psi^T & & & 0\\ & \Psi^T & & \\ & & \ddots & \\ 0 & & & \Psi^T\end{pmatrix}\begin{pmatrix}\underline{f}_1\\ \underline{f}_2\\ \vdots\\ \underline{f}_r\end{pmatrix} \qquad (66)
$$

which is rn equations in $(r+s)m$ unknowns. A necessary condition that (66) has a solution is rank $(A_2) = rn$.* Equation (66) is essentially Equation (43) r times over stacked one over another. For example, if we neglect the intervening zero blocks, the first row of A_4 has the same structure as $\underline{B}$ in (43). It is not hard to convince oneself that if (D,M) is observable, S_r has rank s and A_4 has rank $(s+r)m$, the row dimension of A_4. One way to see this is by partioning Δ and Σ_{ij}.

$$
\Delta = \left(\frac{\Delta_1}{\Delta_2}\right) = \begin{pmatrix} I_m\\ \hline d_1 I_m\\ d_2 I_m\\ \vdots\\ d_s I_m\end{pmatrix}; \qquad \Sigma_{ij} = \left(\frac{\Sigma^1_{ij}}{\Sigma^2_{ij}}\right) = \begin{pmatrix} 0_m\\ \hline s^1_{ij}\, I_m\\ s^2_{ij}\, I_m\\ \vdots\\ s^s_{ij}\, I_m\end{pmatrix}
$$

---

*Again, this is not sufficient for uniqueness. See addendum.

Rearrance the rows of A_4 so that Δ_1 and $\sum_{ij}$ are grouped to the top rows and Δ_2 and $\sum_{ij}{}^2$ are grouped in the lower rows.

$$\widetilde{A}_4 = \begin{array}{cc} \left[\begin{array}{ccc|cccc} \Delta_1 & & O & \sum^1_{11} & \sum^1_{12} & \cdots & \sum^1_{1s} \\ & \Delta_1 \ddots & & \sum^1_{21} & \sum^1_{22} & \cdots & \sum^1_{2s} \\ O & & \Delta_1 & \vdots & \vdots & & \vdots \\ & & & \sum^1_{r1} & \sum^1_{r2} & \cdots & \sum^1_{rs} \\ \Delta_2 & & & \sum^2_{11} & \sum^2_{12} & \cdots & \sum^2_{1s} \\ & \Delta_2 \ddots & & \sum^2_{21} & \sum^2_{22} & \cdots & \sum^2_{2s} \\ & & & \vdots & \vdots & & \vdots \\ & & \Delta_2 & \sum^2_{r1} & \sum^2_{r2} & \cdots & \sum^2_{rs} \end{array}\right] & \begin{array}{c} \\ rm \\ \\ \\ \\ rsm \\ \\ \\ \end{array} \\ \begin{array}{cc} \longleftrightarrow & \longleftrightarrow \\ rm & sm \end{array} & \end{array}$$

$$= \left[\begin{array}{c|c} I_{rm} & 0_{rm \times sm} \\ \hline \begin{matrix} \Delta_2 & & 0 \\ & \Delta_2 \ddots & \\ 0 & & \Delta_2 \end{matrix} & R(S_r{}^T) \times I_m \end{array}\right] \qquad (67)$$

where $R(S_r{}^T)$ is $S_r{}^T$ after row rearrangement, viz.,

$$S_r^{T} = r\!\updownarrow\!\underbrace{\begin{bmatrix} M \\ M(D + d_1 I) \\ M(D^2 + d_1 D + d_2 I) \\ \vdots \\ M(D^{s-1} + d_1 D^{s-2} + \dots + d_{s-1} I) \end{bmatrix}}_{s} \Big\updownarrow sr = \begin{bmatrix} S^1 \\ S^2 \\ S^3 \\ \vdots \\ S^s \end{bmatrix} = \begin{bmatrix} s_{11}^1 & s_{12}^1 & s_{13}^1 & \dots & s_{1s}^1 \\ s_{21}^1 & s_{22}^1 & s_{23}^1 & \dots & s_{2s}^1 \\ \vdots & \vdots & \vdots & & \vdots \\ s_{r1}^1 & s_{r2}^1 & s_{r3}^1 & \dots & s_{rs}^1 \\ s_{11}^2 & s_{12}^2 & s_{13}^2 & \dots & s_{1s}^2 \\ s_{21}^2 & s_{22}^2 & s_{23}^2 & \dots & s_{2s}^2 \\ \vdots & \vdots & \vdots & & \vdots \\ s_{r1}^2 & s_{r2}^2 & s_{r3}^2 & \dots & s_{rs}^2 \\ & & \vdots & & \\ s_{11}^s & s_{12}^s & s_{13}^s & \dots & s_{1s}^s \\ s_{21}^s & s_{22}^s & s_{23}^s & \dots & s_{2s}^s \\ \vdots & \vdots & \vdots & & \vdots \\ s_{r1}^s & s_{r2}^s & s_{r3}^s & \dots & s_{rs}^s \end{bmatrix} \tag{68}$$

$R(S_r^T)$ picks out corresponding rows from each block of S_r^T and groups them into blocks so that the elements of each column in a block have identical subscripts.

a linear transformation of $\mathcal{O}_{D,M}$, the same being true for $R(S_r^T)$ of S_r^T, rank $(R(S_r^T))$ = rank (S_r^T) = s iff (D,M) is observable. Iff this is true, rank $(R(S_r^T) \otimes I_m)$ = ms. We have just shown that $\tilde{A}_4$ and hence A_4 has rank (s+r)m iff (D,M) is observable. Recall the proof of Theorem 1 assumes a general RHS for Equation (43). We will also assume a general RHS for Equation (66). In this case, a solution to (66) exists iff rank (A_2) = nr. To determine the rank of A_2, we need the following lemma (C. T. Chen, page 33, Theorem 2-6).

__Lemma.__ For $A_2 = A_3A_4$,

$$\text{rank } (A_2) = \text{rank } (A_4) - d \tag{71}$$

d is the dimension of the intersection of $R(A_4)$, the range space of A_4, and $N(A_3)$, the null space of A_3.

__Theorem 2__ An asymptotic estimate for a multiple output plant exists if (D,M) is observable, $s \geq \nu-1$ and s and d satisfy $(s+r)m-d = nr$. Given the sufficient conditions are satisfied, the corresponding necessary condition is $s \geq \frac{n-m}{m} r$.

__Proof: sufficient condition__ Given (D,M) is observable, we have just proved rank $(A_4) = (s+r)m$. Given $s \geq \nu-1$, we know that rank (A_2) = nr. A sufficient condition for solution to (66) to exist is rand (A_2) = nr. But by (71), rank (A_2) = rank $(A_4) - d = (s+r)m - d$. Equating the two, $(s+r)m - d = nr$. The condition $s \geq \nu-1$ is required explicitly when r = 1. In that case, A_4 (called __B__ in (43)) spans all of $E^{(s+1)m}$ and d = dim $(N(A_3))$ = n - rank (A_3), (A_3 is called $\underline{A}$ in (43)) = n-n = 0 where $s \geq \nu-1$ ensures rank $(A_3) = n$. When r=1, $(s+r)m - d = nr$ becomes $(s+1)m = n$ or $s = \frac{n-m}{m}$. If $s = \frac{n-m}{m}$ is a sufficient condition, $s \geq \frac{n-m}{m}$ is certainly also sufficient, i.e. adding dimension to the estimator can never hurt us. Since $\frac{n-m}{m}$ is always a lower bound to $\nu-1$, $s \geq \nu-1$ can replace $(s+r)m - d = nr$ when $r = 1$, i.e. $s \geq \nu-1$ can replace $s = \frac{n-m}{m}$. Hence, the sufficient condition of Theorem 2 when r = 1 reduces to that of Theorem 1.

<u>Necessary condition.</u> The necessary condition for (66) to have a solution is

$$\text{rank } (A_2) \leq \min(\text{rank } (A_3), \text{ rank } (A_4)).$$

Given (D,M) is observable, rank $(A_4) = (s+r)m$. Given $s \geq \nu-1$, rank $(A_3) = nr$.

Consider two cases: 1. $nr \leq (s+r)m$

2. $nr > (s+r)m$

1. $\min(nr, (s+r)m) = nr$. Since A_2 has nr independent equations because the sufficient conditions are satisfied, the necessary condition is always satisfied.

2. $\min(nr, (s+r)m) = (s+r)m$. For the necessary condition to be true, rank $(A_2) \leq (s+r)m$. But rank $(A_2) = nr$ because the sufficient conditions are satisfied. $nr \leq (s+r)m$ contradicts the assumption of case 2. Hence 2 can never satisfy the necessary condition.

Rearrange the assumption of case 1, $s \geq \frac{n-m}{m} r$

$(s+r)m - d = nr \Rightarrow s = \frac{(n-m)r+d}{m}$. Since adding dimension to the estimator can never hurt us, $s \geq \frac{(n-m)r+d)}{m}$ might be needed when $(n-m)r+d$ is not an exact multiple of m. Also, $d \geq 0$ and $s \geq \frac{(n-m)r+d}{m}$ renders the necessary condition $s \geq \frac{n-m}{m} r$

redundant. Hence, the necessary and sufficient conditions for an asymptotic estimate to a multiple output plant to exist are (D,M) observable and s $s \geq \max\left(\frac{(n-m)r+d}{m}, \nu-1\right)$.

Before we continue this line of thought further, let us see what is the lower bound to s when we use Theorem 1 (r=1) r times, each for one control input. The way to do so is shown as follows.

$$\dot{x} = Ax + Bu$$

$$B = (b_1|b_2|\ldots|b_r) \tag{72}$$

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_r \end{pmatrix}, \quad u_i,\ i = 1,\ldots, r \text{ are scalar inputs.}$$

$$\dot{x} = Ax + b_1u_1 + b_2u_2 +\ldots+ b_ru_r \tag{73}$$

Each b_iu_i will be considered individually as if it is the only control. A dynamic system

$$t_i \updownarrow \quad z_i = D_iz_i + E_iy + G_iu_i \tag{74}$$

will be constructed so that if $\hat{u} = \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_r \end{pmatrix} \rightarrow Fx$ asymptotically is desired,

$F = \begin{pmatrix} \underline{f}_1^T \\ \underline{f}_2^T \\ \vdots \\ \underline{f}_r^T \end{pmatrix}$, the Luenberger condition (See (26)) and the constraint (See (27))

for the ith control is

$$T_i A - D_i T_i = E_i C \tag{75}$$

$$\underline{h}_i^T C + \underline{m}_i^T T_i = \underline{f}_i^T \tag{76}$$

This problem was solved in deriving Theorem 1. The result is

$$S_i^T \triangleq \begin{bmatrix} \underline{m}_i^T \\ \underline{m}_i^T (D_i + d_1^i I) \\ \underline{m}_i^T (D_i + d_1^i D_i + d_2^i I) \\ \vdots \\ \underline{m}_i^T (D_i^{t_i - 1} + d_1^i D^{t_i - 2} + \dots + d_{t_{i-1}}^i I) \end{bmatrix} \tag{77}$$

where d_j^i, $j = 1, \dots, t_i$ are given by

$$\det (I - D_i) = \lambda^{t_i} + d_1^i \lambda^{t_i - 1} + \dots + d_{t_i}^i = 0 \tag{78}$$

If we denote the t_i rows of E_i by $\underline{e}_1^{iT}$, $\underline{e}_2^{iT}, \dots, \underline{e}_{t_i}^{iT}$, $\underline{h}_i$ and the rows of E_i can be solved from

$$n \updownarrow \overbrace{[A^{T^{t_i}} C^T \vdots A^{T^{t_i-1}} C^T \vdots \cdots \vdots C^T]}^{(t_i+1)m} \underbrace{\begin{bmatrix} I_m & 0_{m \times t_i m} \\ d_1^i I_m & \\ d_2^i I_m & \\ \vdots & S_i^T \otimes I_m \\ d_{t_i}^i I_m & \end{bmatrix}}_{(t_i+1)m} \begin{pmatrix} \underline{h}_i \\ \underline{e}_1^i \\ \underline{e}_2^i \\ \vdots \\ \underline{e}_{t_i}^i \end{pmatrix}$$

$$= (A^{T^{t_i}} + d_1^i A^{T^{t_i-1}} + \ldots + d_{t_i}^i I)\, \underline{f}_i \tag{79}$$

and as long as A and D_i have no eigenvalues in common, T_i is given by

$$T_i = [E_i C A^{ti-1} + (D_i + d_1^i I) E_i C A^{ti-2} + \ldots + (D_i^{ti-1} + d_1^i D_i^{ti-2} + \ldots$$

$$+ d_{t_{i-1}}^i I) E_i C]\; [A^{t_i} + d_1^i A^{ti-1} + \ldots + d_{t_i}^i I]^{-1} \tag{80}$$

The necessary and sufficient condition for (79) to have a solution is $(D_i, \underline{m}_i^T)$ observable and $t_i \geq \max(\frac{n-m}{m}, \nu-1) = \nu-1$ since $\frac{n-m}{m}$ is always a lower bound to $\nu-1$. But for $i = 1, 2, \ldots, r$, $\frac{n-m}{m}$ and $\nu-1$ do not change. Hence, $t_i = t \geq \max(\frac{n-m}{m}, \nu-1) = \nu-1$, i.e., the subscript i of t can be removed. After the above is done for r times, each for a different i, we can combine the results as follows.

By stacking, (75) for $i = 1,2,\ldots, r$ can be written as

$$\underbrace{\begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_r \end{pmatrix}}_{T} A - \underbrace{\begin{pmatrix} D_1 & & 0 \\ & D_2 & \\ 0 & & \ddots \; D_r \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_r \end{pmatrix}}_{T} = \underbrace{\begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_r \end{pmatrix}}_{E} C \tag{81}$$

where we have defined **T**, D and E as the corresponding augmentations. By stacking, (76) for $i = 1,2,\ldots, r$ can be written as

$$\underbrace{\begin{pmatrix} \underline{h}_1^T \\ \underline{h}_2^T \\ \vdots \\ \underline{h}_r^T \end{pmatrix}}_{H} C + \underbrace{\begin{pmatrix} \underline{m}_1^T & & 0 \\ & \underline{m}_2^T & \\ 0 & & \ddots \; \underline{m}_r^T \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_r \end{pmatrix}}_{T} = \underbrace{\begin{pmatrix} \underline{f}_1^T \\ \underline{f}_2^T \\ \vdots \\ \underline{f}_r^T \end{pmatrix}}_{F} \tag{82}$$

where we have defined H and M. Similarly, (74) can be written in totality as

$$\begin{array}{c} t \updownarrow \\ \\ s = rt \updownarrow \end{array} \underbrace{\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_r \end{pmatrix}}_{Z} = \underbrace{\begin{pmatrix} D_1 & & 0 \\ & D_2 & \\ 0 & \ddots & D_r \end{pmatrix}}_{D} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_r \end{pmatrix} + \underbrace{\begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_r \end{pmatrix}}_{E} y + \underbrace{\begin{pmatrix} T_1 b_1 & & 0 \\ & T_2 b_2 & \\ 0 & \ddots & T_r b_r \end{pmatrix}}_{G} \underbrace{\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_r \end{pmatrix}}_{u}$$

(83)

By constructing $u_i \to f_i^T x$ separately, we indeed have accomplished $\hat{u} \to Fx$, as the above three equations show. Since $t \geq \nu-1$ for each subsystem, $s = rt \geq r(\nu-1)$ which is the lower bound to s we sought. Depending on whether $\frac{(n-m)r+d}{m}$ or $r(\nu-1) \geq r\frac{n-m}{m}$ is smaller, we will use Theorem 2 or Theorem 1 r times in order to minimize s. In case of a tie, the computational simplicity of the latter will tip the balance.

Some simple mathematics helps us see when Theorem 2 or Theorem 1 applied r times is useful. Our objective of considering the below minimum-order observer is to realize an estimator of liner functions of x, with dimension $<$ n-m. From Theorem 2, if $\frac{(n-m)r+d}{m} > \nu-1$, the LHS is a lower bound to the dimension s of the below minimum order estimator. But this lower bound cannot be less than n-m if $r > m$! The same applies to Theorem 1 applied r times. Therefore, while Theorem 1 (r=1) looks attractive, its extension to $r > 1$ is useless when $r > m$. Note that even Theorem 1 is useless when $m = 1$ because $\nu-1 \geq \frac{n-m}{m} = n-m$ and we might be better off using the Luenberger observer. Since Theorem 2 assumes $r > 1$, the above discussion $\Longrightarrow 1 < r < m$ has to be satisfied for Theorem 2 to be useful. There is a restriction on n too, namely $n \geq 6$. This is because the minimum r and m that satisfies $1 < r < m$ is $r = 2$, $m = 3$. Assuming $d = 0$, and using the notation [a] = the smallest integer greater than or equal to a, $[\frac{(n-m)r+d)}{m}] = [\frac{(n-3)2}{3}] < n-m = n-3$ iff $n \geq 6$. Most text-book problems do not have $n \geq 6$. An example of a 2-D vehicle on a flexible (elevated) 2-D guideway, the former "driven" by road roughness and wind gust disturbances is studied below: $n = 9$, $m = 4$, $r = 2$. Conditions for $d = 0$ are derived. Simple mathematics shows $n-m = 5$, $[\frac{n-m}{m} r] = 3$, and for that particular problem, $\nu = 3$ $(\nu-1)r = 4$ These will be the dimension of the estimator for Luenberger observer, Theorem 2 and Theorem 1 applied $r = 2$ times respectively.

Let us study the nature of the intersection of $N(A_3)$ with $R(A_4)$. $d = \dim [N(A_3) \cap R(A_4)]$. $d = 0$ when $N(A_3) \perp R(A_4)$. But the row space of A_3, $\underline{R}(A_3)$, $\perp$ $N(A_3)$ by definition of null space. $\therefore$ $\underline{R}(A_3) \perp \perp R(A_4)$ or $\underline{R}(A_3) \subseteq R(A_4)$ when $d = 0$. The maximum dimension of $R(A_3)$ and $R(A_4)$ are nr and $(s+r)m$ respectively. Since $nr \leq (s+r)m$ is required by the necessary condition, this accounts for the direction of "$\subseteq$" above.

We will study the component blocks of A_3 and A_4 in order to discover exactly when $d = 0$. Recall from (66) that

$$\overset{r(s+1)m}{\longleftrightarrow} \quad nr \updownarrow A_3 = \underbrace{\begin{pmatrix} \Gamma & & 0 \\ & \Gamma & \\ 0 & & \ddots \; \Gamma \end{pmatrix}}_{r \text{ blocks}} ; \; \Gamma = [A^{T^s} C^T \vdots A^{T^{s-1}} C^T \vdots \dots \vdots C^T]$$

$$\overset{(r+s)m}{\longleftrightarrow} \quad r(s+1)m \updownarrow A_4 = \left[\begin{array}{ccc:cccc} \Delta & & 0 & \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1s} \\ & \Delta & & \Sigma_{21} & & & \vdots \\ 0 & & \ddots \; \Delta & \vdots & & & \vdots \\ & & & \Sigma_{r1} & & & \Sigma_{rs} \end{array}\right] ; \; \Delta = \begin{bmatrix} I_m \\ d_1 I_m \\ d_2 I_m \\ \vdots \\ d_s I_m \end{bmatrix} \tag{84}$$

(the first three block columns of A_4: r blocks)

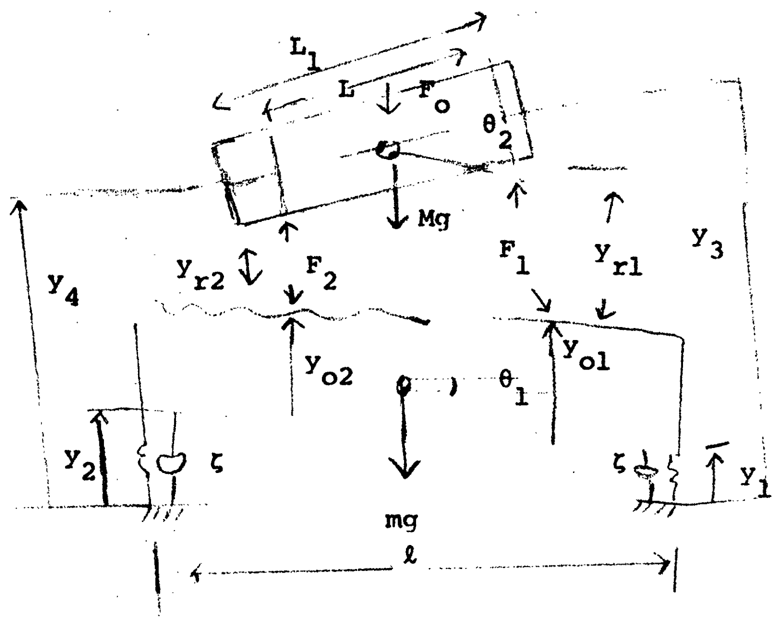
Σ_{ij} given by (65) and (68).

For the first block row of A_3, $(\Gamma | 0_{n \times (r-1)(s+1)m})$, to be in the column space of A_4, $R(A_4)$, we only to have make sure rows of Γ are in the column space of $(\Delta | \Sigma_{11} \; \Sigma_{12} \cdots \Sigma_{1s})$, the first row block of A_4 with the zero blocks after Δ removed. For example, if $\Gamma = (1\ 1)$, the first block row of A_3 is $(1\ 1\ 0\ 0)$. If A_4 is $\left[\begin{array}{cc|cc} \Delta & 0_m & \Sigma_{11} & \Sigma_{12} \\ 0_m & \Delta & \Sigma_{21} & \Sigma_{22} \end{array}\right]$, it is clear that $(1\ 1\ 0\ 0)$ cannot lie in the column space of $\begin{pmatrix} 0_m \\ \Delta \end{pmatrix}$. It is also clear that Σ_{21} and Σ_{22} of $\begin{pmatrix} \Sigma_{11} \\ \Sigma_{21} \end{pmatrix}$ and $\begin{pmatrix} \Sigma_{12} \\ \Sigma_{22} \end{pmatrix}$ may take any values without affecting the latter two's intersection with $(1\ 1\ 0\ 0)$ due to the double zeros. Hence, if $\Gamma = (1\ 1)$ lies in the column space of $(\Delta | \Sigma_{11} \; \Sigma_{12})$, $(1\ 1\ 0\ 0)$ will lie in the column space of A_4.

By similar reasoning, we also require the second block row of A_3, $(0_{n \times (s+1)m} | \Gamma | 0_{n \times (r-2)(s+1)m})$ to lie in the column space of $(\Delta | \Sigma_{21} \; \Sigma_{22} \cdots \Sigma_{2s})$, and so on to the rth block row of A_3. Note that when taken together, these <u>do not</u> mean the rows of Γ lie in the column space of $(\Delta | \Sigma_{11} \; \Sigma_{12} \cdots \Sigma_{1s} \vdots \Sigma_{21} \; \Sigma_{22} \cdots \Sigma_{2s} \vdots \cdots \vdots \Sigma_{r1} \; \Sigma_{r2} \cdots \Sigma_{rs})$ will guarantee $\underline{R}(A_3) \subseteq R(A_4)$. To determine if $d = 0$, we cannot use this <u>one</u> "augmented" test, but must use the test $\underline{R}(\Gamma) \subseteq R(\Delta \vdots \Sigma_{i1} \; \Sigma_{i2} \cdots \Sigma_{is})$ r times, each for a different i, and only when <u>all</u> r tests are satisfied will we know that $d = 0$. These tests will be used to determine the conditions under which $d = 0$, in the 9-th order example below.

Specializing to $r = 1$ may help to understand why Theorem 1 always works, i.e. $d \equiv 0$. When $s \geq \nu-1$, there are n independent rows of A_3, each of which is a (s+1)m-vector. When $(D,\underline{m}^T)$ is observable, A_4 has (s+1)m independent columns, each of which is a (s+1)m-vector. Hence, all of the columns of A_4 span the whole $E^{(s+1)m}$ space. Hence the n independent rows of A_3 ($n \leq (s+1)m \Rightarrow$ at least as many independent equations as unknowns) must always be in $R(A_4) \therefore d \equiv 0$.

<u>Section III</u> <u>Example</u> The vehicle is represented by a mass M and a moment of inertia about c.g., I_M. It is a "beam" with two magnets (wheels) near the ends. The magnets are controlled by two current sources, each independent of the other. (r = 2) The vehicle is subjected to wind gusts F_o acting vertically downwards at c.g. and also to road roughness directly below the magnets. The guideway on which the vehicle moves is a similar "beam" supported by springs and dashpots at its ends. From the passenger's viewpoint, the c.g. of the vehicle is always directly above the c.g. of the guideway <u>when the vehicle is travelling forward at V $\neq$ 0.</u> This is admittedly unrealistic but is possibly the best we can do to model the up-and-down and the rotational motion of the guideway, unless we go into distributed systems. The objective is to minimize mean acceleration at the front and back of the vehicle ($\dot{x}_6$ and $\dot{x}_8$) while keeping the average separation between the magnet and the road surface directly below minimum (x_5 and x_7). I.e. minimize $I = \frac{1}{T}\int_0^T [\dot{x}_6^{\,2} + \dot{x}_8^{\,2} + \rho(x_5^{\,2} + x_7^{\,2})]dt$ as $T \to \infty$, ρ is a weighting factor that specifies the tradeoff between acceleration and road separation.



Appropriately non-dimensionalized, and considering incremental quantities only (i.e. instead of say y_4, consider Δy_4 so that the dead-weight Mg, taken care of by (y_4) nominal, will not have to be considered), x_1 to x_9 will correspond to the physical variables y_1, $\dot{y}_1$, y_2, $\dot{y}_2$, y_{r1}, $\dot{y}_3$, y_{r2}, $\dot{y}_4$, F_o. Defining $\lambda_o = L/\ell$, $\lambda_1 = (L/L_1)^2$, $\mu = M/m$ and neglecting the (small) angles F_1 and F_2 made with the vertical, the state equations are

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{bmatrix} = \left[\begin{array}{c|c|c|c|c|c|c|c|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1-6\lambda_o & -2\zeta(1+6\;_o) & -1+6\lambda_o & -2\zeta(1-6\lambda_o) & a_1{*}\mu(1+6\lambda_o) & 0 & a_3{*}\mu(1-6\lambda_o) & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1+6\lambda_o & -2\zeta(1-6\lambda_o) & -1-6\lambda_o & -2\zeta(1+6\lambda_o) & a_1^{*}\,\mu(1-6\lambda_o) & 0 & a_3{*}\mu(1+6\lambda_o) & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -a_1{*}(1+6\lambda_1) & & -a_3{*}(1-6\lambda_1) & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -a_1{*}(1-6\lambda_1) & 0 & -a_3{*}(1+6\lambda_1) & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & N \end{array}\right] x$$

plus $\longrightarrow$

$$
\begin{bmatrix} 0 & 0 \\ a_2\,\mu(-1-6\lambda_o) & a_4\,\mu(-1+6\lambda_o) \\ 0 & 0 \\ a_2^*\mu(-1+6\lambda_o) & a_4\,\mu(-1-6\lambda_o) \\ 0 & 0 \\ a_2^*(1+6\lambda_1) & a_4\,(1-6\lambda_1) \\ 0 & 0 \\ a_2^*(1-6\lambda_1) & a_4^*(1+6\lambda_1) \\ 0 & 0 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\dot{y}^*_{01} \\ 0 \\ -\dot{y}^*_{o2} \\ 0 \\ n_2 \end{bmatrix}
$$

wind and road-roughness disturbance.

$$
\begin{aligned}
&E(\dot{y}^*_{o1}) = 0, \quad E(\dot{y}^*_{02}) = 0 \\
&E(n_2) = 0 \\
&E(n_2\dot{y}^*_{o1}) = 0 \\
&E(n_2\dot{y}^*_{o2}) = 0 \\
&E\begin{pmatrix} \dot{y}^*_{o1}(t) \\ \dot{y}^*_{o2}(t) \end{pmatrix} (\dot{y}^*_{o1}(\tau) \;\; \dot{y}^*_{o2}(\tau)) \\
&= N_1 \begin{pmatrix} \delta(t-\tau) & \delta(t-\tau+\frac{L}{V}) \\ \delta(t-\tau+\frac{L}{V}) & \delta(t-\tau) \end{pmatrix} \\
&E(n_2(t)n_2(\tau)) = N_2\delta(t-\tau)
\end{aligned}
\tag{85}
$$

Note that road-roughness velocity is modelled as white noise correlated perfectly with a lag $\frac{L}{V}$, the time required to traverse the separation of the magnets at vehicle speed V. Wind-gusts are modelled as first order filtered white noise. N, N_1 and N_2 defines the magnitude of these disturbances. In the above, the magnets are assumed to satisfy the force-separation-current relations: $F_1 = -a_1y_{r1} + a_2u_1$, $F_2 = -a_3y_{r2} + a_4u_2$. The asterik * denotes a non-dimensionalized quantity. The measurements on board the vehicle are assumed to be y_{r1}, $\dot{y}_3$, y_{r2} and $\dot{y}_4$. Accelerations $\ddot{y}_3$ and $\ddot{y}_4$ can be easily measured and then integrated; y_{r1} and y_{r2} not so easily. $\therefore$ m = 4 and y = Cx gives

$$C = \left(\begin{array}{cccc|cccc|c} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right) \tag{86}$$

The problem is to design ρ, μ, λ_o, λ_1 given N, N_1, N_2, V so that I is minimized <u>and</u> frequency response, transient response, and other "classical quantities" are satisfactory. The Wiener filter is used so that $I = I(\rho, \mu, \lambda_o, \lambda_1)$ is minimized <u>and</u> since we are in the frequency domain, can also examine how those responses change with ρ, μ, λ_o, λ_1 and even some ratio between the magnitudes of the disturbances. Suppose this was done and ρ, μ, λ_o, λ_1 are fixed. We now turn to the Kalman filter and its extensions to <u>realize</u> the Wiener filter. Theorem 2 will be used. With $\mu = \frac{1}{2}$, $\zeta = .02$, $\lambda_o = \frac{1}{2}$, $\lambda_1 = .6$, $a_1{}^* = a_3{}^* = 13.3046$, $a_2{}^* = a_4{}^* = 13.3046$, $N = .1$, A becomes

$$A = \left[\begin{array}{c|c|c|c|c|c|c|c|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -4 & -.16 & 2 & .08 & 26.6092 & 0 & -13.3046 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & .08 & -4 & -.16 & -13.3046 & 0 & 26.6092 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -53.2184 & 0 & 26.6092 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 26.6092 & 0 & -53.2184 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .1 \end{array}\right]$$

It was worked out that $\nu = 3$ for this combination of A and C. Since (D,M)

observable is essential to Theorem 2 and we know that for any s, one can always find such a pair. Without loss of generality, (D,M) can take on Luenberger's observable canonical form for multivariate systems.** (See C. T. Chen, pp. 292-295). We have worked out that $[\frac{n-m}{m} r] = 3 > \nu-1$. Hence s = 3 is tried and as the following shows, can work. For s = 3, the Luenberger canonical observation form for D and M are

$$D = \left(\begin{array}{cc|c} 0 & -d_2 & b_2 \\ 1 & -d_1 & b_1 \\ \hline 0 & b_3 & -d_3 \end{array}\right) \quad ; \qquad M = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ \hline 0 & 0 & 1 \end{array}\right) \tag{87}$$

where d_1, d_2, d_3 satisfies $(\lambda^2 + d_1\lambda + d_2)(\lambda + d_3) = 0$ and b_1, b_2, b_3, together with the d_i's, are to be determined. Our immediate objective, however, is to see under what conditions d = 0. Using (87) in (68), we find that

$$S_r^T = \left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 1 & 0 & b_1 \\ 0 & b_3 & d_1-d_3 \\ \hline 0 & b_1b_3 & b_2-b_1d_3 \\ b_3 & -b_3d_3 & b_1b_3-d_1d_3+d_3^2+d_2 \end{array}\right] \tag{88}$$

From (64) and (65), we can construct the component blocks of A_4 in (66). The result is

---

** This requires a demonstration that the BMO observer-based compensator is realization-invariant, however. For the MO case (s = n-m) see Blanvillain and Johnson, Proc. CDC'76 (to appear).

$$[\Delta|\Sigma_{11}|\Sigma_{12}|\Sigma_{13}] = \left[\begin{array}{c|c|c|c} I_4 & 0_4 & 0_4 & 0_4 \\ d_1I_4 & 0_4 & I_4 & 0_4 \\ d_2I_4 & I_4 & 0_4 & b_1I_4 \\ d_3I_4 & 0_4 & b_1b_3I_4 & (b_2-b_1d_3)I_4 \end{array}\right] \tag{89}$$

and

$$[\Delta|\Sigma_{21}|\Sigma_{22}|\Sigma_{23}] = \left[\begin{array}{c|c|c|c} I_4 & 0_4 & 0_4 & 0_4 \\ d_1I_4 & 0_4 & 0_4 & I_4 \\ d_2I_4 & 0_4 & b_3I_4 & (d_1-d_3)I_4 \\ d_3I_4 & b_3I_4 & -b_3d_3I_4 & (b_1b_3-d_1d_3+d_3^2+d_2)I_4 \end{array}\right] \tag{90}$$

Recall that $\underline{R}(\Gamma)$ must lie in the column space of (89) and of (90) for d = 0. One way to guarantee this, without considering what Γ actually is, is to make sure (89), (90) each spans the complete 16-dimensional space, 16 being the dimension of a column vector of (89), (90). For (89) to span E^{16}, one can show that $b_2-b_1d_3 \neq 0$ is enough. (so that the second and the fourth column blocks will be independent.) Siminarly, for (90) to span E^{16}, $b_3 \neq 0$ is enough. Our immediate objective is achieved: $b_2-b_1d_3 \neq 0$ and $b_3 \neq 0$ are sufficient for d = 0 and hence s = 3 will be the minimum order realizable using Theorem 2. To bring the problem to a logical end, we turn to find the optimum b_i's and d_i's, subjected to the above constraints, such that ΔI, the incremental cost due to incomplete state feedback, is minimized. The method by Dakbe-Chen-Powell-Fletcher is one way to do this. Hence, through time-domain results, the Wiener filter is realized. Due to cost separation, the realization is a sequel to the optimal design of ρ, μ, λ_o, λ_1.

In fact, as far as determining the conditions under which $d = 0$ in the above manner is concerned, the problem posed (vehicle on guideway) is relevant only to the extent of specifying n, m, r and ν. For $n \geq 6$, $1 < r < m$, $s = [\frac{n-m}{m} r] > \nu-1$, construct D, M as in (87) with $[\frac{s}{r}]$ partitions, each partition except the one in the lowest right hand side has a dimension r, the last partition has a dimension $s-([\frac{s}{r}]-1)r$. Using MACSYMA, we can generate (88) for all combinations of $n \geq 6$, $1 < r < m$. Conditions for (89), (90), etc. to span the whole $E^{(s+1)m}$ can then be generated and are appropriate to all problems with the same n, m, r combinations.

Section IV

In "An approach to dynamic compensator design for pole assignment", by H. Seragi, Int. J. of Cont., 1975, Vol. 21, No. 6, pp. 955-966, a frequency domain result of Theorem 1 is given, in the context of pole assignment. A summary of that paper is given below.

Seragi considered 1. $r = 1$, $m > 1$

2. $m = 1$, $4 > 1$

but did not give a result that is the counterpart of Theorem 2 where $r > 1$, $m > 1$. Seragi's 1. can be split into three cases: $s \gtreqless \frac{n-m}{m}$. For $\geq$, complete pole assignment is possible, for $<$, only $(n+s)-((s+1)m+s) = n-(s+1)m$ poles can be assigned or all $(n+s)$ poles can be <u>approximately</u> assigned. In Theorem 1, $\geq$ is necessary for the asymptotic estimator to exist, $<$ implies more independent equations than unknowns.

Single-input multi-output systems, $r = 1,\ m > 1$

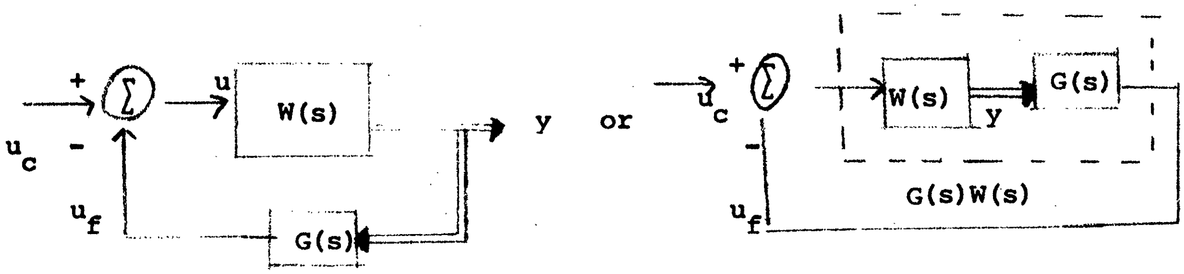
$$x = Ax + bu$$
$$y = Cx$$
$$Y(s) = C(sI-A)^{-1}\,bU(s) = W(s)\ U(s)$$
$$= \frac{w(s)}{F(s)}\,U(s)$$

$$w(s) = \begin{pmatrix} m_{n1}s^{n-1} + \ldots + m_{11} \\ \vdots \\ m_{nm}s^{n-1} + \ldots + m_{1m} \end{pmatrix}, \quad F(s) = s^n + d_n s^{n-1} + \ldots + d_1$$

Let the feedback compensator be $G(s) = \frac{N(s)}{\Delta(s)}$, order s.

$$N(s) = (N_1(s), N_2(s), \ldots, N_m(s)), \quad \Delta(s) = S^s + a_s S^{s-1} + \ldots + a_1$$

$$N_i(s) = b_{si}S^s + b_{s-1,i}S^{s-1} + \ldots + b_{oi}, \ i = 1, \ldots, m.$$



$$\therefore\ \frac{U_f(s)}{U_c(s)} = \frac{G(s)W(s)}{1 + G(s)W(s)}\ ; \quad \frac{Y(s)}{U_c(s)} = \frac{W(s)}{1 + G(s)W(s)} = \frac{w(s)\Delta(s)}{F(s)\Delta(s) + N(s)\cdot w(s)}$$

Let $H(s) = F(s)\Delta(s) + N(s) \cdot w(s)$ = characterist polynominal of the (n+s)th - order closed-loop system

$$= F(s)\Delta(s) + w_1(s)N_1(s) + \ldots + w_m(s)N_m(s)$$

$H(s)$ contains s a_i's and $m(s+1)$ b_i's.

If the desired closed-loop poles are $\lambda_1, \lambda_2, \ldots \lambda_{n+s}$, $H_d(s) = \prod_{i=1}^{n+s} (s-\lambda_i) = s^{n+s} + \hat{d}_{n+s} s^{n+s-1} + \ldots + \hat{d}_1$. Equating coefficients of $H(s)$ and $H_d(s)$, we have

$$Ec = f \tag{91}$$

where E is an $(n+s) \times (s(m+1)+s)$ matrix and f is an $(n+s)$-column vector and c is the $(s + m(s+1))$-column vector of the unknown parameters of the compensator.

$$\text{i.e.} \qquad c = [a_1, \ldots, a_s, b_{o1}, \ldots, b_{s1}, \ldots, b_{om}, \ldots, b_{sm}]^T$$

If $s \geq \frac{n-m}{m}$ and E does not have full rank, increase s until E has full rank. Then solve for c, uniquely when $s = \frac{n-m}{m}$, non-uniquely when $s > \frac{n-m}{m}$. In Theorem 1, $s \geq \max(\frac{n-m}{m}, \nu-1) = \nu-1$ is necessary and sufficient. We conjecture that E will be full rank when $s = \nu-1$.

When $s < \frac{n-m}{m}$ and we can assume in general that E is full rank, E is now $p \times (s(m+1)+s)$, $p < n+s$, we can only specify p closed-loop poles. If we choose to specify <u>all</u> (n+s) poles <u>approximately</u>, we will minimize $||Ec-f||^2$, i.e. $c = (E^TE)^{-1}E^Tf$. The relevance of this result to Theorem 1 is as follows: if $s < \frac{n-m}{m}$, we can build an "approximate" asymptotic estimator using the least-swuare fit solution. Obviously, this applies to Theorem 2 too, when $s < \frac{n-m}{m} r$.

Section V

Explicit solution of the discrete version of Lyaponov Matrix Equation

$A^T PA - P = - Q$

A is an n x n system matrix, P is an n x n sym. matrix that we want to solve for, Q is an n x n sym. "cost" matrix. Define the n(n+1)/2 column vectors composed of rows of the upper triangular P and Q; mathematically,

$$p' = (p_{11}\ p_{12}\ p_{13} \cdots p_{1n}\ p_{22}\ p_{23} \cdots p_{2n}\ p_{33} \cdots p_{(n-1)n}\ p_{nn})$$

$$q_d' = (q_{11}\ q_{12}\ q_{13} \cdots q_{1n}\ q_{22}\ q_{23} \cdots q_{2n}\ q_{33} \ldots q_{(n-1)n}\ q_{nn})$$

subscript d for discrete, (the continuous version defines q differently). We show that the above matrix equation can be re-written as

$$(U_d - I)p = - q_d$$

U_d is an $\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}$ matrix made up of elements of A and I is an $\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}$ identity matrix. U_d is to be constructed as follows.

Label the columns of U_d by 2 indices L and M and the rows of U_d by 2 indices IJ. These labels are identical (in sequence) to the subscripts in the definition of p and q_d. For the columns where L = M, the entry is the single term $a_{LI}\ a_{MJ}$. For all other columns L ≠ M, the entry is the 2-term $a_{LI}\ a_{MJ} + a_{MI}\ a_{LJ}$. Simple, isn't it?

	$L = M$	<u>LM</u>	$L \neq M$
<u>IJ</u>	$a_{LI}\ a_{MJ}$		$a_{LI}\ a_{MJ} + a_{MI}\ a_{LJ}$

LM

IJ	11	12	13	...	22	23	...	33	...	nn
11	a_{11}^2	$a_{11}a_{21}$ $+a_{21}a_{11}$	$a_{11}a_{31}$ $+a_{31}a_{11}$		a_{21}^2	$a_{21}a_{31}$ $+a_{31}a_{21}$		a_{31}^2		a_{n1}^2
12	$a_{11}a_{12}$	$a_{11}a_{22}$ $+a_{21}a_{12}$	$a_{11}a_{32}$ $+a_{31}a_{12}$		$a_{21}a_{22}$	$a_{21}a_{32}$ $+a_{31}a_{22}$		$a_{31}a_{32}$		$a_{n1}a_{n2}$
13										
⋮										
22	a_{12}^2	$a_{12}a_{22}$ $+a_{22}a_{12}$	$a_{12}a_{32}$ $+a_{32}a_{12}$			$a_{22}a_{32}$ $+a_{32}a_{22}$		a_{32}^2		a_{n2}^2
23										
⋮										
33										
⋮										
nn	a_{1n}^2	$a_{1n}a_{2n}$ $+a_{2n}a_{1n}$	$a_{1n}a_{3n}$ $+a_{3n}a_{1n}$			$a_{2n}a_{3n}$ $+a_{3n}a_{2n}$		a_{3n}^2		a_{nn}^2

$\therefore$ $p = -(U_d - I)^{-1} q_d$ is unique (and the inverse exists) iff A is stable. This is just the Lyaponov stability theorem for the rewritten version of the discrete Lyaponov.

The way to see the construction of U_d is a straightforward although tedious rewriting of $A^T PA$ in subscript notation. Since this matrix is syn., only the upper triangular part is used in constructing U_d. This is explained below. Write P in row form and A in column form as follows

$$P = \begin{bmatrix} \underline{p}_1 \\ \underline{p}_2 \\ \vdots \\ \underline{p}_n \end{bmatrix} , \qquad A = \begin{bmatrix} \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \end{bmatrix}$$

$$\therefore (PA)_{ij} = \underline{p}_i \, \underline{a}_j \qquad i = 1, \ldots, n$$
$$j = 1, \ldots, n$$

$$\therefore \quad A^T PA = \begin{pmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{n1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{n1} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{n3} \\ \vdots & & & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \underline{p}_1\underline{a}_1 & \underline{p}_1\underline{a}_1 & \cdots & \underline{p}_1\underline{a}_n \\ \underline{p}_2\underline{a}_1 & \underline{p}_2\underline{a}_1 & & \\ \vdots & & & \\ \underline{p}_n\underline{a}_1 & \underline{p}_n\underline{a}_2 & \cdots & \underline{p}_n\underline{a}_n \end{pmatrix}$$

$$\therefore \quad (A^T PA)_{ij} = \sum_{k=1}^{n} a_{ki} \, \underline{p}_k \, \underline{a}_j , \qquad i = 1, \ldots, n$$
$$j = 1, \ldots, n$$

As noted above, we are only interested in $j \geq i$, the upper $\Delta A^T PA$. There are $\frac{n(n+1)}{2}$ elements in this $\underbrace{\text{upper } \Delta A^T PA.}$ We will take the ij-th <u>element</u> of it to form the IJ-th <u>row</u> of U_d. Note the translation from <u>element</u> to <u>row</u>. This is true by <u>construction</u>. Now comes the difficult part. How do we use $\sum_{k=1}^{n} a_{ki} \, \underline{p}_k \, \underline{a}_j$ to construct the $\frac{n(n+1)}{n}$ <u>elements</u> of the IJ-th <u>row</u> of U_d? You see why LM is labelled like the subscript of p? Because we only want the a_{ki}, $\underline{a}_j$ of $\sum_{k=1}^{n} a_{ki} \, \underline{p}_k \, \underline{a}_j$ in U_d and $\underline{p}_k$ outside U_d :

$$
\begin{array}{c|ccccccc|}
 & & & LM & & & & \\
 & 11 & 12 & 13 \; \ldots & 22 & 23 \; \ldots & & \\
IJ & & & U_d & & & & \\
 & & & & & & &
\end{array}
\begin{pmatrix} \\ p \\ \\ \end{pmatrix}
\qquad
p = \begin{pmatrix} p_{11} \\ p_{12} \\ p_{13} \\ \vdots \\ p_{22} \\ p_{23} \\ \vdots \end{pmatrix}
$$

Without too much ado about the details, the construction of putting "$a_{LI}\, a_{MJ} + a_{MI}\, a_{LJ}$ for $L \neq M$ and $a_{LI}\, a_{MJ}$ for $L = M$" just makes sure the a_{ki}, $\underline{a}_j$ of $\sum_{k=1}^{n} a_{ik}\, \underline{p}_k\, \underline{a}_j$ are put <u>in the right place</u> in U_d so that $U_d^{IJ}\, p = \sum_{k=1}^{n} a_{ki}\, \underline{p}_k\, \underline{a}_j$ where U_d^{IJ} is to denote the IJ-th row of U_d. Now, it is not difficult to see the $-I$ in $(U_d - I)p$ is to account for $-P$ in $A^T PA - P$. Similarly for $-q_d$ and its counterpart $-Q$ in the matrix equation.

<u>Uses of the explicit solution</u>

One application is in the discrete version of T. Johnson's paper on constrained configuration optimal compensator of order s to the linear plant

$$x(i+1) = A\,x(i) + B\,u(i) \quad , \quad x(0) = x_o \qquad (1)$$

$$y(i) = C\,x(i) \quad , \quad \text{rank}\,(C) = m \quad n \qquad (2)$$

S-Compensator

$$\left[\begin{array}{l} z(i+1) = P\,z(i) + N\,y(i) \quad , \quad z(0) = z_o \qquad (3) \\ u(i) = G\,y(i) + H\,z(i) \qquad (4) \end{array}\right.$$

Cost

$$\left[\; J = \sum_{i=0}^{\infty} x^T(i+1)\,Qx(i+1) + u^T(i)\,Ru(i) \right. \qquad (5)$$

parameters to be opt.

$$\left[\; -F \triangleq \begin{pmatrix} G & H \\ N & P \end{pmatrix} \quad , \quad z_o \right.$$

Put (2) in (4) and then in (1),

$$x(i+1) = A\,x(i) + BGC\,x(i) + H\,z(i) \qquad (6)$$

Put (2) in (3),

$$z(i+1) = P\,z(i) + NC\,x(i) \qquad (7)$$

Put (6) & (7) together

$$\begin{pmatrix} x(i+1) \\ z(i+1) \end{pmatrix} = \begin{pmatrix} A+BGC & H \\ NC & P \end{pmatrix} \begin{pmatrix} x(i) \\ z(i) \end{pmatrix} \qquad (8)$$

$$\triangleq \hat{A} \begin{pmatrix} x(i) \\ z(i) \end{pmatrix} = \left(\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} B & 0 \\ 0 & I \end{pmatrix} \underbrace{\begin{pmatrix} G & H \\ N & P \end{pmatrix}}_{-F} \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} \right) \begin{pmatrix} x(i) \\ z(i) \end{pmatrix}$$

Put (1) in (5),

$$J = \sum_{i=0}^{\infty} (Ax(i) + Bu(i))^T Q(Ax(i) + Bu(i)) + u^T(i)Ru(i) \qquad (9)$$

Put (2) in (4) and then in (9) and expand and neglecting the i,

$$J = \sum (x^T A^T + x^T C^T G^T B^T + z^T H^T B^T) \, Q \, (Ax + BGCx + BHz) + (x^T C^T G^T + z^T H^T) R \, (GCX + Hz)$$

$$= \sum_{i=0}^{\infty} \begin{pmatrix} x(i) \\ z(i) \end{pmatrix}^T \left(\begin{array}{c|c} (A + BGC)^T \, Q \, (A + BGC) + (GC)^T \, RGC & (A + BGC)^T \, QBH + (GC)^T \, RH \\ \hline H^T B^T Q \, (A + BGC) + H^T RGC & H^T (B^T QB + R) \, H \end{array} \right) \begin{pmatrix} x(i) \\ z(i) \end{pmatrix}$$

$$\triangleq \sum_{i=0}^{\infty} \begin{pmatrix} x(i) \\ z(i) \end{pmatrix}^T \hat{Q} \begin{pmatrix} x(i) \\ z(i) \end{pmatrix} \qquad (10)$$

From Kalman and Bertram, 1960: "Cont. System Anay. and Design Via the second method of Lyaponov, II Discrete Time Syst.", Trans. ASME June, 1960, p. 394-400, 10 can be written as

$$J = \begin{pmatrix} x(0) \\ z(0) \end{pmatrix}^T \hat{P} \begin{pmatrix} x(0) \\ z(0) \end{pmatrix} = \text{tr } [\hat{P} \, W] \qquad (11)$$

where $\hat{P}$ satisfies

$$\hat{A}^T \hat{P} \hat{A} - \hat{P} = - \hat{Q} \text{ and } W \triangleq \begin{pmatrix} x(0) \\ z(0) \end{pmatrix} \begin{pmatrix} x(0) \\ z(0) \end{pmatrix}^T \qquad (12)$$

where $\hat{A}$ is defined in (8). The matrices $\hat{A}$, $\hat{P}$ and $\hat{Q}$ are $(m+s) \times (n+s)$.

Let $n+s = \alpha$.

Define
$$\hat{p}' = (\hat{p}_{11}\ \hat{p}_{12} \cdots \hat{p}_{1n}\ \hat{p}_{22}\ \hat{p}_{23} \cdots \hat{p}_{\alpha\alpha})$$
$$w' = (w_{11}\ 2w_{12} \cdots 2w_{1n}\ w_{22}\ 2w_{23} \cdots w_{\alpha\alpha})$$
$$\hat{q}' = (\hat{q}_{11}\ \hat{q}_{12} \cdots \hat{q}_{1n}\ \hat{q}_{22}\ \hat{q}_{23} \cdots \hat{q}_{\alpha\alpha})$$

$\therefore$ (11) is $J = w'\hat{p}$ (13)

and (12) is $(U_d - I)\hat{p} = -\hat{q}$ (14)

Differentiate (14) wrt F_{ij}

$$\frac{\partial U_d}{\partial F_{ij}} \hat{p} + (U_d - I) \frac{\partial \hat{p}}{\partial F_{ij}} = - \frac{\partial \hat{q}}{\partial F_{ij}}$$

$$\therefore \frac{\partial \hat{p}}{\partial F_{ij}} = -(U_d - I)^{-1} \left(\frac{\partial U_d}{\partial F_{ij}} \hat{p} + \frac{\partial \hat{q}}{\partial F_{ij}} \right)$$

$$= (U_d - I)^{-1} \left[\frac{\partial U_d}{\partial F_{ij}} (U_d - I)^{-1} \hat{q} - \frac{\partial \hat{q}}{\partial F_{ij}} \right] \qquad (15)$$

where we have used (14)

Similarly, with $z_o' = (z_{o1}\ z_{o2} \cdots z_{os})$

$$\frac{\partial \hat{p}}{z_{ol}} = 0 \text{ since } U_d \neq U_d(z_o) \text{ and } \hat{q} \neq \hat{q}(z_o) \qquad (16)$$

$\therefore$ Differentiating J wrt F_{ij}, (13) becomes

$$\frac{\partial J}{\partial F_{ij}} = \frac{\partial w'}{\partial F_{ij}} \hat{p} + w' \frac{\partial \hat{p}}{\partial F_{ij}}$$

$$= w' \frac{\partial \hat{p}}{\partial F_{ij}} \quad , \quad w \neq w(F_{ij})$$

$$= w'(U_d - I)^{-1} \left[\frac{\partial U_d}{\partial F_{ij}} (U_d - I)^{-1} \hat{q} - \frac{\partial \hat{q}}{\partial F_{ij}} \right] \tag{17}$$

on using 15, 17 = 0 for necessary conditions for opt wrt F_{ij}

Differentiating J wrt z_{oi}, 13 becomes

$$\frac{\partial J}{\partial z_{oi}} = \frac{\partial w'}{\partial z_{oi}} \hat{p} + w' \frac{\partial \hat{p}}{\partial z_{oi}}$$

$$= \frac{\partial w'}{\partial z_{oi}} \hat{p} \quad \text{, due to 16}$$

$$= - \frac{\partial w'}{\partial z_{oi}} (U_d - I)^{-1} \hat{q} \text{ , due to 14} \tag{18}$$

18 = 0 for necessary condition for opt wrt z_{oi}

Without too much modification, we can change 1 to

$$x(i+1) = A\,x(i) + Bu(i) + \underline{G}v(i) \tag{19}$$

where v(i) is white or coloured noise disturbance.

In IEEE Trans. on Auto Control, April 1975, Kurtaran considers a plant 19 and measurement

$$y(i) = C\,x(i) + \underline{H}\,w(i) \tag{20}$$

where w(i) is white noise and proceeds to design a "suboptimal" (with the possibility of y(i) feedforward) compensator of the form 3 and 4

$$E\left[\begin{pmatrix} v(i) \\ w(i) \end{pmatrix} (v(i)' \; w'(i))\right] = \tilde{V}, \qquad \tilde{V} \geq 0$$

The augmented system is

$$\begin{pmatrix} x(i+1) \\ z(i+1) \end{pmatrix} = \hat{A}\begin{pmatrix} x(i) \\ z(i) \end{pmatrix} + \left[\begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} B & 0 \\ 0 & I \end{pmatrix}\underbrace{\begin{pmatrix} G & H \\ N & P \end{pmatrix}}_{-F}\begin{pmatrix} 0 & \underline{H} \\ 0 & 0 \end{pmatrix}\right]\begin{pmatrix} v(i) \\ w(i) \end{pmatrix} \qquad (21)$$

where $\hat{A}$ is as in 8.

With $J \triangleq E(x(i)'\,Q\,x(i) + u(i)'\,R\,u(i)),\; i \to \infty$ (22)

which is different from 5, Kurtaran shows that if

$$E\left[\begin{pmatrix} x(i) \\ z(i) \end{pmatrix} (v'(i) \quad w'(i))\right] = 0$$

$$J = E\left[\begin{pmatrix} x(i) \\ z(i) \end{pmatrix}' (Q + \tilde{C}'\,\tilde{F}'\,\tilde{R}\,\tilde{F}\,\tilde{C}) \begin{pmatrix} x(i) \\ z(i) \end{pmatrix}\right.$$

$$\left. + \begin{pmatrix} v(i) \\ w(i) \end{pmatrix}' \tilde{H}'\,\tilde{F}'\,\tilde{R}\,\tilde{F}\,\tilde{H} \begin{pmatrix} v(i) \\ w(i) \end{pmatrix}\right], \quad i \to \infty \qquad (23)$$

where $\tilde{Q} = \begin{array}{c} \\ \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix} \\ \overset{\leftrightarrow}{n} \quad \overset{\leftrightarrow}{s} \end{array} \begin{array}{l} \updownarrow n \\ \updownarrow s \end{array}$, $\tilde{R} = \begin{array}{c} \\ \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} \\ \overset{\leftrightarrow}{r} \quad \overset{\leftrightarrow}{s} \end{array} \begin{array}{l} \updownarrow r \\ \updownarrow s \end{array}$, $\tilde{H} = \begin{array}{c} \\ \begin{pmatrix} 0 & \underline{H} \\ 0 & 0 \end{pmatrix} \\ \overset{\leftrightarrow}{q_1} \quad \overset{\leftrightarrow}{q_2} \end{array} \begin{array}{l} \updownarrow m \\ \updownarrow s \end{array}$

$v(i) \updownarrow q_1$, $w(i) \updownarrow q_2$, $\tilde{C} = \begin{array}{c} \\ \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} \\ \overset{\leftrightarrow}{n} \quad \overset{\leftrightarrow}{s} \end{array} \begin{array}{l} \updownarrow m \\ \updownarrow s \end{array}$, $\tilde{B} = \begin{array}{c} \\ \begin{pmatrix} B & 0 \\ 0 & I \end{pmatrix} \\ \overset{\leftrightarrow}{r} \quad \overset{\leftrightarrow}{s} \end{array} \begin{array}{l} \updownarrow m \\ \updownarrow s \end{array}$

$u(i) \updownarrow r$ $\qquad y(i) \updownarrow m$

With $\Sigma(i) \triangleq E\left[\begin{pmatrix} x(i) \\ z(i) \end{pmatrix} (x'(i) \;\; z'(i))\right]$, the steady state Σ is

(24) $$\Sigma = \underbrace{(\tilde{A} - \tilde{B} F \tilde{C})}_{\hat{A}} \Sigma (\tilde{A} - \tilde{B} F \tilde{C})' + \underbrace{(\tilde{G} - \tilde{B} F \tilde{H}) \tilde{V} (\tilde{G} - \tilde{B} F \tilde{H})'}_{\tilde{V}} ; \tilde{G} = \begin{array}{c} \overset{\leftrightarrow}{q_1} \quad \overset{\leftrightarrow}{q_2} \\ \begin{pmatrix} \underline{G} & 0 \\ 0 & 0 \end{pmatrix} \end{array} \begin{array}{l} \updownarrow n \\ \updownarrow s \end{array}$$

Writing 23 in terms of Σ & V,

(25) $$J = \operatorname{tr} [\underbrace{(\tilde{Q} + \tilde{C}' F' \tilde{R} F \tilde{C})}_{\bar{Q}} \Sigma] + \operatorname{tr} (\underbrace{\tilde{H}' F' \tilde{R} F \tilde{H}}_{\bar{R}} \tilde{V})$$

$\bar{Q} = \begin{pmatrix} C' G' R G C & C' G R H \\ H' R G C & H' R H \end{pmatrix}$ is different from $\hat{Q}$ in (10) because

(22) is different from (5).

Using the technique previously developed, we can write 24

(26) $$\Sigma = \hat{A} \Sigma \hat{A}' + \bar{V} \qquad [\text{cf } A^T P A - P = -Q]$$

as $\sigma = -(U_d - I)^{-1}\,\bar{v}$

Similarly, 25 is

27 $$J = \bar{q}'\sigma + \tilde{v}\,\bar{r}'$$

$$\frac{\partial J}{\partial F_{ij}} = \frac{\partial \bar{q}'}{\partial F_{ij}}\,\sigma + \bar{q}'\,\frac{\partial \sigma}{\partial F_{ij}} + \tilde{v}\,\frac{\partial \bar{r}'}{\partial F_{ij}}$$

$$= -\frac{\partial \bar{q}'}{\partial F_{ij}}\,(U_d - I)^{-1}\,\bar{v} + \bar{q}'(U_d - I)^{-1}\left[\frac{\partial U_d}{\partial F_{ij}}\,(U_d - I)^{-1}\,\bar{v} - \underbrace{\cancel{\frac{\partial \bar{v}}{\partial F_{ij}}}}_{0}\right]$$

$$+ \tilde{v}\,\frac{\partial \bar{r}'}{\partial F_{ij}} \qquad (28)$$

28 = 0 for necessary condition for opt. wrt F_{ij}.

$$\frac{\partial J}{\partial z_{oi}} = \frac{\partial}{\partial z_{oi}}\,(\bar{q}'\,\sigma + \tilde{v}\,\bar{r}') \qquad (29)$$

$\equiv 0$ since none of the terms in( ) are functions of z_{oi}

This is agreeable since J in 22 is steady-state performance and should be independent of z_o.

Summary of Results

The explicit solution of $A'PA - P = -Q$ is shown to be useful in

1 the "transient" problem (an opt. z_o is required)

$$x(i+1) = A\,x(i) + B\,u(i), \quad x(0) = x_o$$

$$y(i) = C\,x(i)$$

$$z(i+1) = P\,z(i) + N\,y(i), \quad z(0) = z_o$$

$$u(i) = G\,y(i) + H\,z(i)$$

$$J = \sum_{i=0}^{\infty} x(i+1)'\,Q\,x(i+1) + u(i)\,R\,u(i)$$

(2) the "steady-state" problem (any z_o will be optimal)

$$x(i+1) = Ax(i) + Bu(i) + \underline{G}v(i),\ x(0) = x.$$

$$y(i) = Cx(i) + \underline{H}w(i)$$

$$J = E[x(i)'\ Q\ x(i) + u(i)'\ Ru(i)],\ i\to\infty$$

This contains as special case where $\underline{H} \equiv 0$. (If $\underline{G} \equiv 0$, so will J and <u>any</u> stable system is optimal!)

(3) As if life is not complicated enough, the "hybrid case". (an optimal z_o required)

$$x(i+1) = A\ x(i) + B\ u(i) + \underline{G}\ v(i),\ x(0) = x_o$$

$$y(i) = C\ x(i) + \underline{H}\ w(i)$$

$$J = E\left(\sum_{i=0}^{N-1} x'(i+1)\ Q\ x(i+1) + u'(i)\ Ru(i)\right)$$

This contains as special cases where either $\underline{G} \equiv 0$ or $\underline{H} \equiv 0$ or both $\underline{G} \equiv 0$ and $\underline{H} \equiv 0$. When both $\underline{G} \equiv 0$, $\underline{H} \equiv 0$, 3 is the finite-horizon counterpart of 1 and G, H, P, N will <u>not</u> be constant but functions of i. Furthermore, the non-steady-state equation.

(30) $$P(i+1) = A'P(i)A + Q$$

is to replace $P = A'PA + Q$. Whereas the latter is rewritten as

(31) $$(U_d - I)p = -q,$$

(30) can be rewritten similarly as

(32) $$p(i+1) = U_d(p(i) + q$$

The finite-horizon formulation is conceptually and technically not too different from the infinite-horizon formulation. The difference is one of tedium: the former has to find optimal G, H, P and N for every i, $0 \leq i \leq N-1$.

Three more cases can be listed, each corresponds to the previous three with $x(0) = x_o$ random.

4) the "transient" problem, random x_o

Replace J by E(J) and x_o by $E(x_o)$ and $x_o x_o'$ by $E(x_o x_o')$ and $x_o z_o'$ by $E(x_o z_o') = E(x_o) z_o'$

5) the "steady-state" problem, random x_o

No change is needed, random x_o does not affect s.s. result.

6) the "hybrid case", random x_o

Replace x_o, $x_o x_o'$, $x_o z_o'$ by $E(x_o)$, $E(x_o x_o')$ and $E(x_o) z_o'$ respectively.

When an optimal z_o is required, it can be shown that $z_o^* = z_o^*(x_o)$ or $z_o^*(E(x_o))$ in the random x_o case. When x_o is unknown and one does not want to assign an $E(x_o)$ and an $E(x_o x_o')$, one can modify the cost to

a) minimum maximum

$$\text{Min } M, \quad M = \max_{x_o} \frac{J}{x_o^T x_o} = \max \frac{x_o^T P(0)\, x_o}{x_o^T x_o}$$

b) minimum maximum relative to optimal

$$\text{min } L, \quad L = \max_{x_o} \frac{J}{J_{opt}} = \max \frac{x_o^T P(0) x_o}{x_o^T P_{opt}(0) x_o}$$

P opt is solution of the mat. Ricatti's Equation for optimal regulator,

y = x

(c) minimum an upper bound of L, call it B

(d) average ratio

$$J_1 = \underset{x_o}{\text{ave}} \frac{J}{x_o^T x_o} = \underset{x_o}{\text{ave}} \frac{x_o^T P(0)\, x_o}{x_o^T x_o}$$

It can be shown that

$$M = \lambda_{max}(P(0)), \quad \lambda_{max}(\cdot) = \text{maximum eigenvalue}$$

$$L = \lambda_{max}(P_{opt}^{-1}(0)\, P(0))$$

$$B^2 = ||P(0)|| = \sum_{i=1}^{n} \lambda_i^2(P(0)) = \sum_{i,j=1}^{n} p_{ij}^2$$

$$J_1 = tr\,(P(0)) = \sum_{i=1}^{n} p_{ii}$$

All these four criteria give a solution independent of x_o. At present, the use of M and L is limited to e.g. second order plant because they involve finding the maximium eigenvalue of P(0) which is known only in terms of the parameter matrices G, H, P and N. Symbolic manipulation is required to find the eigenvalues of P(0). What is more difficult is to find the maximum eigenvalue when the eigenvalues are expressed symbolically. Then, minimum $\lambda_{max}(G, H, P, N) \Rightarrow \frac{\partial \lambda_{max}(G,H,P,N)}{\partial F_{ij}} = 0$

In view of the above difficulty, B^2 is used. Dabke 1970 (IEEE AC-15, pp. 120-122) shows that $B^2 = \sum_{i,j=1}^{n} p_{ij}^2$ which can be easily implemented by the Dabke-Chen-Shieh method. So can $J_1 = tr\,(P(0)) = \sum_{i=1}^{n} p_{ij}$ which was first used by Kleinman & Athans 1968. More explicitly, it can be shown that

$B^2 = p' \, S \, p$, p defined at the beginning of this section and S is an $[\frac{n(n+1)}{2}]^2$ diagonal matrix with diagonal elements 1 or 2 so that p_{ij} get weighted by 1 and p_{ij}, $i \neq j$ get weighted by 2.

$$\therefore \frac{\partial B^2}{\partial F_{ij}} = \frac{\partial p'}{\partial F_{ij}} Sp + p'S \frac{\partial p}{\partial F_{ij}} = 0$$

Similar, we can show that

$$J_1 = -(S - 2I)p, \text{ I is an } [\frac{n(n+1)}{2}]^2 \text{ identity matrix}$$

$$\therefore \frac{\partial J_1}{\partial F_{ij}} = -(S - 2I) \frac{\partial p}{\partial F_{ij}} = 0$$

Section VI

Extension of C. F. Chen and L. S. Shieh, "A Note on Expanding $PA + A^T P = -Q$", IEEE Trans. Auto Cont., AC-13, pp. 122-123, Feb. 1968.

In the above reference, an algorithm was developed to expand the Lyaponov Equation into a $(n(n+1)/2)^2$ matrix equation which can be solved by matrix inversion. In the present extension, an algorithm for accomplishing the above symbollically is developed. For the incomplete state-feedback problem, a Lyaponov Equation $P(A-BFC) + (A-BFC)'P = -Q-C'F'RFC = -\hat{Q}$ is involved in which F is to be optimized. This necessitates a symbolic expansion.

The algorithm expands $PA + A^T P = -Q$ into $Up = -\frac{1}{2}q$. If A is n x n, p and q are $n(n+1)/2$ dimensional vectors defined as

$$p = (p_{11}, p_{12}, p_{13}, \ldots, p_{1n}\ p_{22}, p_{23}, \ldots, p_{2n}, p_{33}, \ldots, p_{n-1,n-1}, p_{nn})^T$$

$$q = (q_{11}, 2q_{12}, 2q_{13}, \ldots, 2q_{1n}, q_{22}, 2q_{23}, \ldots, 2q_{2n}, q_{33}, \ldots, 2q_{n-1,n-1}, q_{nn})^T$$

and U is an $(n(n+1)/2)^2$ non-singular matrix that is a function of a_{ij}, the elements of A. U will be non-singular if A is stable. The algorithm has to do with writing U in terms of a_{ij}.

1 Label the rows and columns of U by the subscripts of the elements of p. E.g. for a 3 x 3 system, the p vector is $(p_{11}\ p_{12}\ p_{13}\ p_{22}\ p_{23}\ p_{33})^T$. The row and column labels of U will be

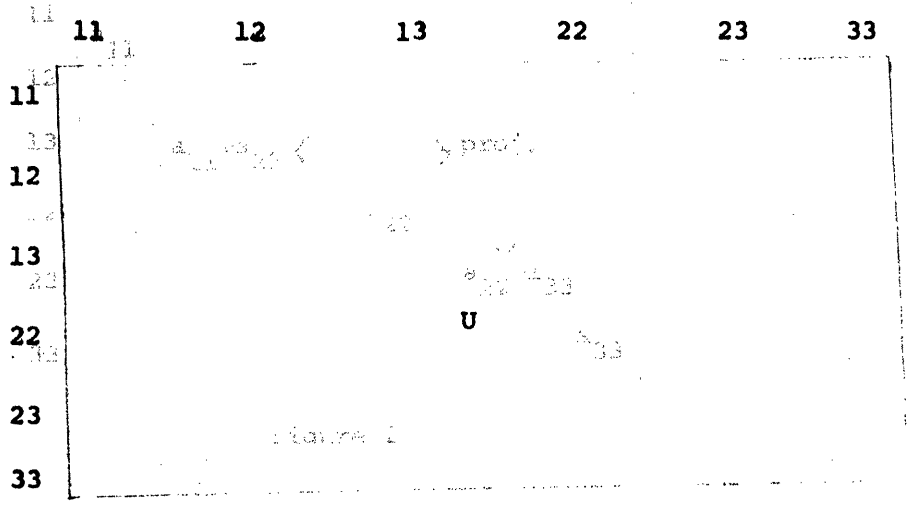


Figure 1

Each element of U is referred to by the quadruple (row label, column label).

(2) Generate diagonal 1 Element (ii, ii) $= a_{ii}$

2 Element (ij, ij) $= a_{ij} + a_{ij}$, $i \neq j$

A diagonal element with 1 term is a singleton, 2 terms a dublet. See Figure 2.

(3) Generate above-diagonal elements.

Each element above the diagonal has 2 "projections" onto the diagonal, a horizontal and a vertical. E.g., element (12,23) projects onto the 2 diagonal elements (12, 12) and (23,23) which are $a_{11} + a_{22}$ and $a_{22} + a_{33}$ respectively. Four cases can be distinguished.

(1) When the 2 projections are singletons, set the element to 0. E.g., Projection of element (11,22) are a_{11} and a_{22}. $\therefore$ Element (11,22) = 0.

$$
\begin{array}{c|cccccc|}
 & 11 & 12 & 13 & 22 & 23 & 33 \\
\hline
11 & a_{11} & & & & & \\
12 & & & & & & \\
13 & & a_{11}+a_{22} & & & \text{proj.} & \\
22 & & & & a_{22} & & \\
23 & & & & & a_{22}+a_{33} & \\
33 & & & & & & a_{33} \\
\hline
\end{array}
$$

Figure 2

(2) When the "a" subscripts of the 2 projections do not intersect, set the element to 0. E.g., Projections of element (13,22) are $a_{11}+a_{33}$ and a_{22}. Since 2 does not "intersect" 1 and 3, element (13,22) = 0.

(3) When the 2 projections are singleton and a dublet, the element takes the "a" subscripts of the dublet and the 2 subscripts of the entry are in decreasing order. E.g. projections of element (11,12) are a_{11} and $a_{11} + a_{22}$. Element (11,12) = a_{21}.

(4) When the 2 projections are 2 dublets, the element has "a" subscripts of the unequal subscripts of the 2 dublets and the 2 subscripts of the element are in decreasing order. E.g. the projections of element (12,13) are $a_{11} + a_{22}$ and $a_{11} + a_{33}$. Element (12,13) = a_{32}.

(4) Generate the below-diagonal elements.

Each element below the diagonal has "a" subscripts in reversed order of the corresponding entry above the diagonal. E.g. element (22,12) = a_{12} because the corresponding entry above the diagonal is element (12,22) = a_{21}.

That's all for the algorithm. It is less complicated than it looks. If the computer takes symbolic inputs and gives symbolic output, this algorithm is not difficult to program.

Application

In the Dakbe method originally developed for output feedback, one needs to find $\frac{\partial U}{\partial F_{ij}}$. This is possible only when U is explicitly written as functions of F_{ij} as the above algorithm provides. The following is a reminder of the context in which the problem arises.

$$\dot{x} = Ax + Bu$$
$$y = Cx$$
$$\dot{z} = Pz + Ny$$
$$u = Gy + Hz$$
$$J = \int_0^\infty (x'Qx + u'Ru)dt$$

Find $F = -\begin{pmatrix} G & H \\ N & P \end{pmatrix}$ such that J is minimized.

After some substitutions, $J = tr(Y\ P)$ where $Y = \begin{pmatrix} x_o \\ z_o \end{pmatrix} (x_o' \ z_o')$ and P satisfies $P(A-BFC) + (A-BFC)'P = Q-C'F'RFC$. Let $y = (y_{11}, 2y_{12}, 2y_{13}, \dots, 2y_{1n}, y_{22}, 2y_{23}, \dots, 2y_{2n}, y_{33}, \dots, 2y_{n-1,n-1}, y_{nn})^T$

$\therefore J = y^T p$. The necessary conditions for a minimized J is $\frac{\partial J}{\partial F_{ij}} = y^T \frac{\partial p}{\partial F_{ij}} = 0$. Using the above algorithm, the Lyaponov equation can be written as $U\ p = -\frac{1}{2}q$. Differentiating wrt F_{ij}, $\frac{\partial U}{\partial F_{ij}} p + U \frac{\partial p}{\partial F_{ij}} = -\frac{1}{2} \frac{\partial q}{\partial F_{ij}}$

$$\frac{\partial p}{\partial F_{ij}} = - U^{-1} \left(\frac{1}{2} \frac{\partial q}{\partial F_{ij}} + \frac{\partial U}{\partial F_{ij}} p \right). \quad \text{Since } p = - \frac{1}{2} U^{-1} q,$$

$$\frac{\partial p}{\partial F_{ij}} = - \frac{1}{2} U^{-1} \left(\frac{\partial q}{\partial F_{ij}} - \frac{\partial U}{\partial F_{ij}} U^{-1} q \right)$$

The necessary conditions are then

$$y^T U^{-1} \left(\frac{\partial U}{\partial F_{ij}} U^{-1} q - \frac{\partial q}{\partial F_{ij}} \right) = 0.$$

When n is small, these conditions (non-linear) can be set up and solved for optimal F_{ij}. When n is big, U^{-1} is not easy. The Powell-Eletcher numerical solution will be useful.

Section VII

Some worked examples, using MACSYMA

E1 Example 1

Output Feedback. 2 x 2 example from Kurtaran and Sidar. Transformed so that $C = (0 \quad I_m)$, $A = \begin{pmatrix} -1 & 0 \\ 1 & -2 \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $C = (0 \ 1)$, $Y = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $K = (1.32811 \quad .15323)$, $Q = \begin{pmatrix} 3 & -2 \\ -2 & 2 \end{pmatrix}$ $F = f$, $A\text{-}BFC = \begin{pmatrix} -1 & -f \\ 1 & -2f-2 \end{pmatrix}$. Using the algorithm just developed, $U = \begin{pmatrix} -1 & 1 & 0 \\ -f & -2f-3 & 1 \\ 0 & -f & -2f-2 \end{pmatrix}$,

$$\frac{\partial U}{\partial f} = \begin{pmatrix} 0 & 0 & 0 \\ -1 & -2 & 0 \\ 0 & -1 & -2 \end{pmatrix} .$$

$$Q + C'F'RFC = \begin{pmatrix} 3 & -2 \\ -2 & 2+3f^2 \end{pmatrix}, \quad q = \begin{pmatrix} 3 \\ -4 \\ 2+3f^2 \end{pmatrix}, \quad \frac{\partial q}{\partial f} = \begin{pmatrix} 0 \\ 0 \\ 6f \end{pmatrix} .$$

$y = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix}$. The necessary condition is $216 f^4 + 936 f^3 + 1257 f^2 + 468 f - 102 = 0$. Of the 4 solutions, 2 are real, 2 are complex; only $f = .1502733$ gives a stable overall system. The optimal $J = .90408136$. The resultant closed-loop poles are at -1.128179376 and -2.172367223. This agrees with Kurtaran and Sidar who used a method involving Lagrange multipliers. We have avoided using the latter by solving the constraints explicitly, helped by the algorithm.

J1 Example 2

Johnson Compensator, 2 x 2 example from Blanvillain. Transformed so that $C = (0 \quad I_m)$, $A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $C = (0 \quad 1)$, $Y = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$,

$$K = (-1 \quad -1), \; Q = \begin{pmatrix} 5 & -1 \\ -1 & 1 \end{pmatrix} \quad F = \begin{pmatrix} g & h \\ n & p \end{pmatrix} \; , \; A - BFC = \begin{pmatrix} -1 & 1 & 0 \\ 1 & g-1 & h \\ 0 & n & 0 \end{pmatrix} .$$

Using the algorithm, we find

$$U = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & g-2 & n & 1 & 0 & 0 \\ 0 & h & p-1 & 0 & 1 & 0 \\ 0 & 1 & 0 & g-1 & n & 0 \\ 0 & 0 & 1 & h & p+g-1 & n \\ 0 & 0 & 0 & 0 & h & p \end{bmatrix}$$

This turns out to be too big for MACSYMA to invert. So $p = -3$ and $n = 4$ are assumed. This involves fixing the compensator but leaves the feedback gains free. Proceeding,

$$\frac{\partial U}{\partial g} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \; ; \; \frac{\partial U}{\partial g} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$q = \begin{bmatrix} 5 \\ -2 \\ 0 \\ 1+g^2 \\ 2gh \\ h^2 \end{bmatrix} \; ; \; \frac{\partial q}{\partial g} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2g \\ 2h \\ 0 \end{bmatrix} \; ; \; \frac{\partial g}{\partial h} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2g \\ 2h \end{bmatrix}$$

$y^T = (1\ 4\ 0\ 5\ 0\ 0)$, assuming $z_o z'_o = 0$. The necessary conditions are rather long polynomials in g and h. Of the 8 pairs of solutions, only $(g,h) = (-1.2464551, -.83837253)$ produces a stable composit system. The optimal cost is 12.1696497. Blainvillain, using a minimum order Luenberger observer (dim = 1) found the optimal cost is 12.071. The discrepancy is due to our fixing $p = -3$ and $n = 4$. Had we used $p = -1$, $n = 4$, the optimal compensator found by Blainvillain, the opt. $(g,h) = (-1.41421363, -.35355337)$ and the opt. cost is 12.071068. This is a numerical "proof" that Johnson's compensator $\equiv$ Luenberger observer when the both have the same dimension. Rom proved this rigorously in his dissertation.

<u>Details of Example 2</u>.

$$Q + C'F'RFC = \begin{pmatrix} 5 & -1 & 0 \\ -1 & 1+g^2 & gh \\ 0 & gh & h^2 \end{pmatrix}$$

When $p = -1$ and $n = 4$, the 2 necessary conditions are

$$\begin{aligned} \frac{\partial}{\partial g}: \quad & -1024h^5 - 256g^2h^4 - 512gh^4 + 1536h^4 - 384g^3h^3 \\ & + 1088g^2h^3 - 1088h^3 - 208g^4h^2 + 1024g^3h^2 \\ & - 1200g^2h^2 - 480gh^2 + 1024h^2 - 48g^5h + 296g^4h \\ & - 504g^3h - 132g^2h + 1024gh - 672h - 4g^6 + 28g^5 \\ & - 57g^4 - 28g^3 + 256g^2 - 336g + 144 = 0 \end{aligned}$$

$$\frac{\partial}{\partial h}: \quad 256gh^4 - 512h^4 + 384g^2h^3 - 1152gh^3 + 768h^3$$
$$+ 208g^3h^2 - 800g^2h^2 + 784gh^2 - 32h^2 + 48g^4h$$
$$- 216g^3h + 184g^2h + 304gh - 384h + 4g^5$$
$$- 20g^4 + g^3 + 142g^2 - 264g + 144 = 0.$$

Solving these polynomials is not as difficult as setting them up which requires U^{-1} symbolically. With 4 symbols in U, this already overloads MACSYMA, and yet this is just a small (n=2, s=1) problem! Using the Fletcher-Powell algorithms symbolic matrix inversion is not necessary.

Section VIII **Recent References**

Relevance Degree

1 Sirisena, H. R. and Choi, S. S., "Optimal pole placement in linear multivariable systems using dynamic o/p flb", Int. J. Control, 1975, vol. 21, No. 4, 661-671.

1 Leondes, C. T. and Yocum, J. F., Jr., "Opt. Observers for Continuous Time Linear Stochastic Systems", Automatic, Vol. 11, pp. 61-73, 1975.

1 Jameson, A. & O'Malley, R. E., Jr., "Cheap Control of Time-Invariant Regulator", Applied Math and Opt., an international journal, Vol. 1, No. 4, 1975, pp. 337ff.

$$J(\varepsilon) = \frac{1}{2}\int_0^\infty (x^T Q x + \varepsilon^2 u^T R u)\,dt, \ \varepsilon \to 0.$$

2 Johnson, C. D., "On observers for systems with unknown and inaccessible inputs", Int. J. Cont., 1975, Vol. 21, No. 5, pp. 825-831.

3 Extension of "Pearson and Ding 1969 and Ferguson and Rekasius 1969", pole placement result but not opt.

(a) Brasch, F. M., Jr. & Pearson J.B. "Pole Placement Using Dynamic Compensators", IEEE AC-15, #1, Feb. 1970, p. 34 ff

(b) Ding, C. Y., Brasch, F. M., Jr., Pearson, J.B., "On Multivariable Linear Systems", same issue, pp. 96 ff.

Recent References Continued

Relevance Degree

c Shmari Reza and Vacroux, A. G., "On pole assignment in linear systems with fixed order compensators", Int. J. Cont., 1973, Vol. 17, No. 2, pp. 397-404.

1 H. Seraji, "An approach to dynamic compensator design for pole assignment", International Journal of Control, Vol. 21, No. 6, June 1975, pp. 955-966.

contains frequency domain results of the 30-odd pages of write-up included. However, does not consider the general case of $r > 1$ and $m > 1$.

1 R. K. Cavin, C. Thisayakorn, and J. W. Howze, "On the Design of Observers with Specified Eigenvalues", IEEE Trans., Aug. 1975, pp. 568-570.

Addendum to Section II:

The results of Section II, though revealing, are not necessary and sufficient. This is because the stated conditions of Theorems 1 and 2 guarantee a solution of (43) and (66), respectively, but not a unique solution. This means, in effect, that s may still be further reduced, using up the additional degrees of freedom. As a result of the cost separation lemma we have that the compensation problem is equivalent to

$$\text{Min} \quad \Delta J = \text{tr}[P\, T \textstyle\sum_{o} T'] \tag{1}$$

subject to

$$TA - DT = EC \tag{2}$$

$$HC + MT = F \tag{3}$$

$$D'P + PD = -M'RM \tag{4}$$

where

D: s x s	A: n x n	F: r x n
P: s x s	E: s x m	
T: s x n	H: r x m	
C: m x n	M: r x s	

In comparing Design Methods #1 and #2 introduced in Section II, it is instructive to compare the numbers of free parameters. For #2, the number of free parameters is the number of elements of D and M (s^2+rs) plus the number of free parameters in H and E, which is less the number of equations in (66) or [(r+s)m-nr] $\geq$ 0 by Theorem 2 the number of elements of H and E. So the total number of free parameters is [(r+s)(m+s)-nr]. Method #1 leaves T, M, and H free, which accounts for [r(m+s)+ns] parameters. It would appear that the method which leaves more free parameters admits the possibility of achieving a lower minimum. We ask, then if

$$s(m+s) - rn \lesseqgtr ns \quad \text{(eliminating common terms)}$$

Consider the example of $n = 10$, $m = r = 3$, for which $s = 7$ satisfies the conditions of Theorem 2 (with equality); we then find

$$(\#2)\ 40 < 70\ (\#1)$$

So that method #1 has more free parameters. For more than a few inputs, similar conclusions can be derived in more general situations. Generally, the trick in solving (1)-(4) is to satisfy as much of (2)-(4) as possible without eliminating too many degrees of freedom. Method #1 can in fact be applied to the BMO design problem, and the essentials are as follows:

Let

$$D = TAF^{\dagger}M \qquad (F^{\dagger}F = I) \tag{5}$$

$$E = TAF^{\dagger}H \tag{6}$$

which imply that (2) is satisfied. Then (4) becomes

$$M'F^{\dagger\prime}A'T'P + PTAF^{\dagger}M = M'RM \tag{7}$$

$$P = M'\tilde{P}M; \quad \tilde{P}: \quad r \times r, \text{ symmetric} \tag{8}$$

so that (7) becomes

$$F'A'T'M'\tilde{P} + \tilde{P}MTAF = -R \tag{9}$$

Thus (1)-(4) is replaced by (1), (3) and (9), with T, M, H to be found. Notice that this problem only involves the product

$$S \doteq MT \quad (t \times n) \tag{10}$$

so that (3) and (9) are replaced by

$$F = HC + S \tag{11}$$

$$F'A'S'\tilde{P} + \tilde{P}SAF = -R \tag{12}$$

But now any S,F may be uniquely written as

$$S = S_1 C + S_2 \text{ with } S_2C' = 0 \tag{13}$$

$$F = F_1 C + F_2 \text{ with } F_2C' = 0 \tag{14}$$

So that (11) implies the two conditions

$$H + S_1 = F_1 \tag{15}$$

$$S_2 = F_2 \qquad (\text{so } S_2 \text{ is fixed!}) \tag{16}$$

and thus we know

$$S = S_1C + F_2 \tag{17}$$

Using this back in the original problem, we have the far simpler parameterization of the problem:

$$\min_{S_1} \Delta J = \mathrm{tr}[\tilde{P}(S_1 C \textstyle\sum_o C'S_1' + S_1 C \sum_o F_2' + F_2 \sum_o C'S_1' + F_2 \sum_o F_2')] \tag{18}$$

subject to

$$(F^{\dagger\prime}A'F_2' + F^{\dagger\prime}A'C'S_1')\tilde{P} + \tilde{P}(S_1CAF^{\dagger} + F_2AF^{\dagger}) = -R \tag{19}$$

with S_1: r x m of a very elegant dimension! Obviously, (19) is a type of projection of (4) which takes into account the control and observation properties of the original problem. Eq. (18) and (19) contain the essence of the BMO compensation problem; more complete results will be reported in a future publication.