Control and Aggregation (I)

Sung-Shin Han*

Utilization of the aggregation concept applied in economics has been a traditional way of describing the state of an economic system and of predicting the future economic conditions. In addition, certain aggregate variables have also played a crucial role as indicators of the business cycle. Quick examples would be the price index, the productivity changes, the industrial production index, GNP, and so on. The methods of aggregation could be either simple summations, like GNP, or sophisticated weighted average, like the price index. Aggregation could be applied to individuals, as in personal income, to commodities, as in price index, or over both of them, as in consumption expenuditre. Furthermore, monthly, quarterly, semiannual, and annual data determine the degree of aggregation over time. The size of the econometric mcdel is dependent upon the degree of aggregation of both individuals and commodities and the length of the lag structure is determined by the degree of aggregation over time[5]. The order (dimension) of the problem of analyzing the model's dynamic properties and of optimally controlling it is an increasing function of its size and its length of lag structure of the model.

In general, the justification invoked for the use of aggregates is to secure information on the micro-level, which cannot be analyzed without aggregation due to formidable computational difficulties or unavailability of micro data.

As far as aggregation from the micro variables to the macro variables is concerned, the methods of aggregation are based on a priori knowledge with

*Assistant Professor, Department of Econcmics, Yonsei University. This is a part of the author's dissertation submitted to the University of Pennsylvania, 1977.

theoretical considerations. If a priori knowledge asserts that some variables always move together, they may be aggregated into a single variable, which will be an appropriately weighted average of the original variables, as in the Hicks-Lange criterion[11]. With relation to this criterion, Ando-Simon[2] analyzed the aggregation of variables in the so-called nearly completely decomposable dynamic system. This system can be approximated by a single variable with respect to each subsystem as if the whole system were completely decomposable. In fact, it is impossible to derive, in a way which is free of contradictions, an aggregated macro-relationship of the usual type from given micro economic theories. As Ando-Simon claimed, the conditions for exact aggregation are very severe. Hence, whether these conditions are strictly satisfied in any practical situation is not really important since any model we employ is no more than an approximate description of reality. The aggregation problem should be dealt with out of the sheer necessity and based on practicality rather than validity. Therefore, as far as aggregation is concerned, we are looking for heuristic rules and criteria which will yield satisfactory approximations under certain conditions.

For instance, we can consider the following sequence of findings in relation to the identification and estimation of the structural macro-econometric model. Suppose that the restrictions on the structural parameters for identification in a macro-econometric model are based on the micro structural model from micro-economic theory. Theil's finding[17] of a relationship between the macro and micro structural model indicates that the macro parameters are weighted sums of the parameters of the underlying structural micro relations. The sum of weights is 1 in the case of correspondence, 0 in case of noncorrespondence, and the weights are not independent of the micro parameters to be weighted, and depend in general on all parameters of the entire micro system.

A subset of the miro-model, which does not include at least one variable included in the other subset, is to be aggregated into one macro equation. Then the macro-parameters of the excluded variables are weighted sums of all the micro-parameters with the sum of weights being 0(which are supposed to receive zero-restrictions in the macro-model). Clearly the sum of weights being zero does not imply that the weighted sum is zero. We might assume that the weighted sum is close to zero under the certain conditions. Fisher[7] on this approximate restriction for the identification has shown that, if these restrictions are approximately satisfied, the estimates of the structural parameters, obtained by the k-class estimators, are asymptotically consistent in the limit.

So far we have discussed a justification of the usual practice in macro-econometric model building. The conditions of the approximation can hardly be checked in reality, but are simply assumed to be satisfied. We mentioned before that the dimension of the model is determined by the degree of aggregation. Even with such aggreation, many models are too large to be estimated by the simultaneous estimation methods like 3SLS and FIML. Therefore, they are usually estimated by OLS, 2SLS, or instrumental variable estimator, because of the computational difficulties and degree of freedom problems. The same problem happens in the control of a large-scale model in the sense of computational difficulties. Even though a mathematical programming approach might be useful in the case of short-time horizon, the number of instruments in such an approach increases in proportion to the time horizon[6, 10].

Therefore, this chapter discusses the further aggregation of the given model in order to reduce the dimensions of the model to manipulable size. Ando-Simon[2] did not fail to indicate this possibility and termed it a multi-level hierarchical aggregation. Thus we can consider the construction of the macro-econometric model as an aggregation by the first-level hierarchy. Now we shall discuss second level hierarchy in this paper. For simple analysis without confusing the main idea, the first level aggregation is assumed to be exactly true without aggregation error. As in the case of micro-macro relationship, there is no absolute criterion except for a special case such as the complete block diagonal coefficient matrix. But the heuristic criteria are to be investigated. This paper considers the deterministic econometric model without error terms, for the stochastic model with additive error terms and the deterministic model do not make any difference in the sense of control rule by the certainty equivalence principle. However, this paper does not consider the econometric model with the stochastic coefficients.

This paper consists of six sections. The first section describes the canonical representation of a dynamic econometric model, which would be a good criterion for satisfactory aggregation and would play a role similar to principal component analysis in estimation theory. Section 2 defines an aggregation matrix for the later convenient analysis and demonstrates the dynamic behavior of an aggregated model to derive the propagation of errors due to aggregation. Section 3 investigates several possible aggregations of the model, with examples in the case where the aggregation matrix can be determined a priori. Section 4 demonstrates the approximate determination of an aggregation matrix using a priori knowledge, as in Ando-Simon's near decomposability. Section 5 analyzes the role of the aggregative model in control and the effect of the control rule of the aggregative model on the disaggregative model as a stabilization policy. Section 6 summarizes the results briefly and suggests future research areas for control of the large-scale models.

1. Canonical Representation of Dynamic Econometric Model

Let the econometric model given in the state-variable form,

$$(1.1)\qquad \underset{n\times 1}{\boldsymbol{y}_t} = \underset{n\times n}{\boldsymbol{A}}\ \underset{n\times 1}{\boldsymbol{y}_{t-1}} + \underset{n\times m}{\boldsymbol{B}}\ \underset{m\times 1}{\boldsymbol{x}_t}$$

where $\boldsymbol{y}_t$ and $\boldsymbol{x}_t$ are the endogenous variables and the instruments, respectively. Two kinds of canonical representations are discussed, (1) on the endogenous variable space, and (2) on the instrument space. The former is useful for the next three sections and the latter for the later sections. We begin by describing a canonical representation in the endogenous variable space.

Without loss of generality, suppose the matrix $\boldsymbol{A}$ has n distinct characteristic roots without multiplicity. Then

(1. 2) $$A=P\Lambda P^{-1}$$

where Λ and P are diagonal matrix and the matrix of the characteristic vectors. Another useful form of equation (1. 2) from simple matrix manipulation is

(1. 3) $$A=\sum_{i=1}^{n} \lambda_i p_i q_i$$

where $P=(p_1, p_2,\cdots,p_i,\cdots,p_n)$ for column vector p_i

$$P^{-1}=Q=\begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_i \\ \vdots \\ q_n \end{bmatrix} \text{ for row vector } q_i$$

and λ_i is the ith diagonal element of the diagonal matrix Λ. Let

(1. 4) $$\underset{n\times 1}{Z_t} = \underset{n\times n}{P^{-1}} \ \underset{n\times 1}{y_t}$$

Then equation (1. 1) can be rewritten as

(1. 5) $$Z_t=\Lambda Z_{t-1}+Dx_t$$

where $D=P^{-1}B$

Now let us consider the second kind of canonical form on the instrument space[12]. To derive this, it is assumed that the condition of controllability holds, or

(1. 6) $$\underset{n\times(nm)}{\Gamma} = \{B, AB, A^2B,\cdots,A^{n-1}B\} \text{ has rank } n.$$

Furthermore, controllability index v is defined as the smallest positive integer for which the matrix

(1. 7) $$\underset{n\times(mv)}{\Gamma_v} = \{B, AB, A^2B,\cdots,A^{v-1}B\}$$

has rank n, such that $v\leq n$. Without loss of generality, the column vectors b_1 of B are linearly independent and $m\leq n$. Then it is possible to define matrix S of the form,

(1. 8) $S=\{b_1, Ab_1,\cdots,A^{v_1-1}b_1, b_2, Ab_2,\cdots,A^{v_2-1}b_2,\cdots,A^{v_m-1}b_m\}$ which has rank n

such that $v_i \leq v$ (See Luenberger[12] for the proof). By the change of the columns,

(1.9) $$S=\{b_1, b_2, \cdots, b_m, Ab_1, Ab_2, \cdots, A^{v_m-1}b_m\}$$

As in the first case, taking the similarity transformation does not change the dynamic property of the econometric model. We can transform in the following manner,

(1.10) $$Z_t=S^{-1}y_t$$

Then

$$Z_t=\tilde{A}Z_{t-1}+\tilde{B}x_t$$

where $\tilde{A}=S^{-1}AS$

and $\tilde{B}=S^{-1}B$

Let

(1.12) $S^{-1}=\begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$ where e_i is a row vector

Then

(1.13) $$\tilde{B}=S^{-1}B=\overset{m}{\begin{bmatrix} I \\ \cdots \\ 0 \end{bmatrix}}\begin{matrix} m \\ \\ n-m \end{matrix}$$

because $S^{-1}S=I$ and the matrix S is a nonsingular square matrix.

2. Aggregation

Before considering dynamic aggregation methods, the general algebra on the aggregation matrix (which is defined in this section), and the dynamic properties of the aggregated model should be discussed. The last section, in some sense, considers an exact aggregation. In the other words, the exact aggregation implies that there exists a unique inverse transformation by an equivalence transformation which does not change the whole set of characteristic roots or the dynamic behavior of the model. As for aggregation in the usual meaning, equivalence does not hold, but dynamic behavior can be approximated to a certain degree for the purpose of computational and analytical convenience. For this purpose the aggregation matrix does not have to be a square matrix like the matrix S or P in the canonical representation, but could rather be a rectangular matrix whose number of columns is larger than that of rows. The original model could, thus,be reduced to an aggregative model with less dimension.

(Definition) Aggregation of the dynamic model can be defined as the transformation of the n-dimensional endogenous variable vector into the l-dimensional vector such that l is less than n.[1]

Thus we can define the aggregation matrix C as

$$\underset{l\times 1}{Z_t} = \underset{l\times n}{C}\ \underset{n\times 1}{y_t} \qquad \text{where } l<n. \tag{2.1}$$

Suppose the dynamics of the Z_t-vector be constructed as

$$Z_t = FZ_{t-1} + DX_t \tag{2.2}$$

The premultiplying C in equation(1.1),

$$Cy_t = CAy_{t-1} + CBx_t \tag{2.3}$$

Comparing equations (2.2) and (2.3), the following equation should hold,

$$CA = FC \tag{2.4}$$

$$CB = D \tag{2.5}$$

To investigate the dynamic behavior of the equation (2.2), the following theorem on the matrix equation (2.4) would be useful.

(Theorem 1) The matrix A and F of equation (2.4) have common characteristic roots if and only if there exists a nontrivial aggregation matrix C.

(Proof) (necessity)[9]

$$\text{Let } A = U\Lambda_A U^{-1} \text{ and } F = V\Lambda_F V^{-1}$$

where Λ_A and Λ_F are diagonal. Then

$$CU\Lambda_A U^{-1} = V\Lambda_F V^{-1}C, \text{ or } V^{-1}CU\Lambda_F = \Lambda_F V^{-1}CU.$$

Then the equation,

$$\tilde{C}\Lambda_A - \Lambda_F\tilde{C}, \text{ where } \tilde{C} = V^{-1}CU,$$

defines $l\times n$ linear equation system, or $(\lambda_{Ai} - \lambda_{Fj})C_{ji} = 0$

If $\lambda_{Ai} \neq \lambda_{Fj}$ for all i and j, $\tilde{C}_{ji} = 0$.

1) This definition is similar to that of minimal-order observer. See; Aoki, M., *Optimization of Stochastic Systems*, Academic Press, 1967 (pp. 250-265).

Therefore $C=V\tilde{C}U^{-1}$ must be equal to zero **matrix.**

If $\lambda_{Ai}=\lambda_{Fj}$ for some i,j, $\tilde{C}_{ji}$ could be assigned an arbitrary non-zero number. With such construction, C will be non-zero. As an alternative proof, let us consider the characteristic polynomials,

$|\lambda I-A|=\phi_A(\lambda)$: n-degree

and $|\lambda I-F|=\phi_F(\lambda)$: l-degree

where $|\cdot|$ denotes determinant.

If there is no common root λ, $\phi_A(\lambda)$ will generate n-dimensional subspace I_n of R^{n+l}, and $\phi_F(\lambda)$ l-dimensional subspace I_l, of R^{n+l}. Furthermore I_n and I_l have the common null vector, so that $R^{n+l}=I_n+I_l$. Therefore I_n and I_l are mutually independent subspaces, so there does not exist any common vector except for the null vector.

(Sufficiency) Suppose the characteristic roots of A be $\lambda_1,\cdots,\lambda_n$ and the corresponding characteristic vectors $u_1,u_2,\cdots,u_n$. Choose $Cu_1\neq 0$ and there necessarily exists such u_i.

Then $CAu_i=C\lambda_i Iu_i$

$=\lambda_i Cu_i=FCu_i$

Therefore $(F-\lambda_i)Cu_i=0$ q.e.d.

Furthermore we have to note that, if the matrix C consists of l independent row vectors, the matrices A and F have exactly the l common characteristic roots. To study the validity of the aggregation, dynamic behavior of the aggregation error will be discussed.

Let the aggregation error be

(2.6) $e_t=Z_t-Cy_t$.

Then $e_t=FZ_{t-1}+Gx_t-CAy_{t-1}-CBx_t$

$=F(e_{t-1}+Cy_{t-1})-CAy_{t-1}$

$=Fe_{t-1}+(FC-CA)y_{t-1}$

Therefore the solution form of e_t will be

(2.7) $e_t=F^te_0+\sum_{\tau=1}^{t}F^{t-\tau}(FC-CA)y_{\tau-1}$

In the macro-model building, it cannot be expected that $FC=CA$, because matrix A of the micro-model is unknown and coefficient matrix F is estimated statistically. Under these circumstances, we need a constant adjustment procedure, even though the aggregation variable z is well defined at the initial stage, i.e., $e_0=0$, because of the second term in equation(2.7). Furthermore, if the underlying model is unstable, the prediction bias will blow up because of y_{t-1} in equation(2.7). Therefore, when the aggregated model is used for the purpose of forecasting, we have to be very careful. Even if the underlying model is stable, long-run prediction would not be recommended except in a special case such as Ando-Simon's nearly complete decomposability condition of the matrix A [2]. (This point will be discussed later in detail.) As far as stabilization policy is concerned, the situation is a little bit different. Suppose the planner uses the aggregated model (2.2) for the long-run policy and compute the feedback equation of Chapter II.

(2.8)
$$\begin{aligned} x_t &= -GZ_{t-1} - g_t \\ &= -GCy_{t-1} - g_t \end{aligned}$$

The existence of stabilization policy depends upon the characteristic roots of $(A-BGC)$ where the matrix G is a function of matrices F and D. However, the error, e_t, will affect g_t which is a function of the linear quadratic tracking equation. Therefore as in the case of forecasting with the aggregative model, we need a fine-tuning procedure on the vector, g_t, of the linear feedback equation (2.8) at every point of time. This can be done by recalculation of the linear quadratic tracking equation, only if the maximum absolute eigenvalues of $(A-BGC)$ is less than 1. (This point will be investigated in Section 5 in detail.)

3. Some Examples of Aggregation

The construction of the aggregated model for computational convenience and analysis depends upon the matrices C and F. If matrix F is given explicity, matrix C could be derived from the Theorem 1. But in reality, matrix

F is not known and is to be constructed to have some desired qualitative aspects which directly depend upon the construction of the aggregation matrix C. Thus this section studies the construction of matrix F with some qualitative aspects which will be determined by the planner. For the purpose of illustration, three examples will be taken: (1) possession of the subset of characteristic roots of the original model, (2) possession of the largest moduli of characteristic roots, and (3) possession of the subset of characteristic roots with a restriction on the representation of endogenous variables.

(Example 1) Suppose the row vectors, C_i of matrix C are mutually orthogonal. This is the case of the usual definition of the aggregated variables in macro-economics like durable consumption, nondurable consumption, plant-equipment investment, housing investment and so on, which are the summation of each group without overlap. Let

(3.1) $$C_iC_j'=0 \quad \text{if } i\neq j$$
$$=R_i \quad \text{if } i=j$$

Suppose the matrix F have the following form,

(3.2) $$F=\sum_{i=1}^{l}\lambda_i\hat{p}_i\hat{q}_i$$

where $\hat{p}_i$ and $\hat{q}_i$ are l-dimensional column and row vector and λ_i is a subset of characteristic roots of the original model in equation (1.3). Then p_i and q_i can be constructed from equation (1.3) as

(3.3) $$\hat{p}_{ij}=C_jp_i$$

(3.4) $$\hat{q}_{ji}=q_iC_j'/R_j \qquad i,j=1,\cdots,l$$

where $$\hat{p}_i=(\hat{p}_{i1},\cdots,\hat{p}_{ij},\cdots,\hat{p}_{kl})'$$

and $$\hat{q}_{ji}=(\hat{q}_{1i},\cdots,\hat{q}_{ji},\cdots,\hat{q}_{li})$$

The above can be derived as

(3.5) $$F=CAC'(CC')^{-1}$$

when $C_j\cdot p_i=0$ where $j=1,\cdots,l$ and $i=l+1,\cdots,n$

(Example 2) Without loss of generality, the diagonal matrix, Λ, of the equation (1.2) assumes to be arranged in the order of the largest characteristic root in the absolute value.

Defining

$$C=\begin{bmatrix} q_1 \\ \vdots \\ q_l \end{bmatrix} \tag{3.6}$$

we can construct the aggregated model as

$$Z_{it}=\lambda_i Z_{it-1}+Dx_t \tag{3.7}$$

where $i=1,\cdots,l.$

By the same reasoning as example 1, we can construct the C matrix such that $C\perp(p_{l+1},\cdots,p_n)$ where '$\perp$' denotes 'orthogonal to', and the matrix F can be computed by equation (3.5), though, in this case, the coefficient matrix of lagged aggregated endogenous variables is not diagonal as in equation (3.7).

(Example 3)[4]

Aggregation via characteristic vectors may have the subset of characteristic roots of the original model, but the basic problem is that the aggregated variable cannot be interpreted in economic meaning. Thus this example considers the partitioning of the endogenous variables to have economic meaning. In other words, partitioning the endogenous vector as

$$y_t=\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} \begin{matrix} l \\ n-l \end{matrix} \tag{3.8}$$

we want to construct the model as

$$y_{1t}=Fy_{1,t-1}+Dx_t \tag{3.9}$$

with the characteristic roots of F being a subset of characteristic roots of matrix A. Clearly this is a further restriction of the example 1 or example 2. Equation (3.9) might not be suitable to the purpose of forecasting because it depends upon the contribution of the omitted variables, y_{2t} in the equation (3.9). In the case of policy analysis, accurate calculation of the dynamic multiplier is most important. If we could approximate the dynamic multiplier with fairly good accuracy, policy analysis could be carried out approximately. This example tries to approximate the dynamic multiplier with the similar technique of example 1 and example 2. In equation(1.1), the dynamic multiplier would be governed by the following equation,

$$(3.10)\qquad \Delta y_t = A\Delta y_{t-1} + B\Delta x_t,$$

$$\Delta x_t = x_t - \tilde{x}_t, \text{ and } \Delta y_t = y_t - \tilde{y}_t$$

where '~' denotes the nominal path and Δy_t indicates the increment due to Δx_t.

Then the solution form will be

$$(3.11)\qquad \Delta y_t = \sum_{r=\tau}^{t} A^{t-\tau} B\Delta x.$$

Without loss of generality, let $\Delta x_\tau = 1$ for simplicity and define

$$(3.12)\qquad M = \sum_{\tau=1}^{t} A^{t-\tau} B.$$

Equation (3.12) can be rewritten from equation (1.3) as

$$(3.13)\qquad M = \sum_{\tau=1}^{t} \left(\sum_{i=1}^{n} \lambda_i^{t-\tau} p_i q_i\right) B$$

where p_i and q_i are normalized as $p_i' p_i = q_i q_i' = 1$. Changing the summation sign in equation (3.13),

$$(3.14)\qquad M = \sum_{i=1}^{n} \left(\sum_{\tau=1}^{t} \lambda_i^{t-\tau}\right) p_i q_i B$$

$$= \sum_{i=1}^{n} \frac{\lambda_i^t - 1}{\lambda_i - 1} p_i q_i B$$

To get the form of equation (3.9) with the first l characteristic roots, equation (3.14) should have the form as

$$(3.15)\qquad \text{where}\quad \tilde{M} = \sum_{i=1}^{l} \frac{\lambda_i^t - 1}{\lambda_i - 1} \begin{bmatrix} p_{1i} \\ \vdots \\ p_{li} \end{bmatrix} (q_{i1}, \cdots, q_{il}) \begin{bmatrix} b_1 \\ \vdots \\ b_l \end{bmatrix}$$

$$p_i' = (p_{1i}, \cdots, p_{li}, \cdots, p_{ni})',$$

$$q_i = (q_{i1}, \cdots, q_{il}, \cdots, q_{in}),$$

and

$$B' = (b_1, \cdots, b_l, \cdots, b_n)',$$

$$b_i : m \times 1 \text{ row vector}$$

The matrix representation of equation (3.15) will be

$$(3.16)\qquad \tilde{M} = P_{11}\ \Lambda^* Q_{11} B_1$$

where

$$P = \left[\begin{array}{c|c} P_{11} & P_{12} \\ \hline P_{21} & P_{22} \end{array}\right] \begin{array}{l} l \\ n-l \end{array}$$

$$\quad l \qquad n-l$$

$$Q = P^{-1} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{matrix} l \\ n-l \end{matrix}$$
$$\qquad\quad l \quad n-l$$

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \begin{matrix} l \\ n-l \end{matrix}$$
$$\quad m$$

and

$$\Lambda^* = \begin{bmatrix} \frac{\lambda^t{}_1 - 1}{\lambda_1 - 1} & 0 & 0 \\ 0 & \ddots \frac{\lambda^t{}_i - 1}{\lambda_i - 1} & 0 \\ 0 & 0 & \ddots \frac{\lambda_l{}^t - 1}{\lambda_l - 1} \end{bmatrix}$$

To derive the form of equation (3. 10) from the approximated dynamic multiplier matrix, $\tilde{M}$, equation (3. 16) can be rewritten, using the similar transformation technique, as

(3. 17)
$$\tilde{M} = P_{11}\Lambda^* P_{11}^{-1} P_{11} Q_{11} B_1$$
$$= \sum_{\tau=1}^{t} F^{t-\tau} D$$

where
$$F = P_{11}\Lambda_1 P_{11}^{-1}$$
$$D = P_{11} Q_{11} B_1,$$

and

$$\Lambda_1 = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots \lambda_i & 0 \\ 0 & \cdots & \ddots \lambda_l \end{bmatrix}$$

Then we can construct the dynamic multiplier equation like equation(3. 10) as

(3. 18)
$$\Delta y_{1t} = F \Delta y_{1t-1} + D \Delta x_t$$

In the more convenient form of matrix F, the following matrix identity will be useful.

(Theorem 2)

$$F = P_{11}\Lambda_1 P_{11}^{-1}$$
$$= A_{11} + A_{12} P_{21} P_{11}^{-1}$$

where

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{matrix} l \\ n-l \end{matrix}$$
$$\quad l \quad n-l$$

(Proof)

$$\begin{aligned} P_{11}\Lambda_1P_{11}^{-1} &= P_{11}\Lambda_1(Q_{11}P_{11}+Q_{12}P_{21})P_{11}^{-1} \\ &\quad +P_{12}\Lambda_2(Q_{21}P_{11}+Q_{22}P_{21})P_{11}^{-1} \\ &\qquad (\because QP=I) \\ &= P_{11}\Lambda_1Q_{11}+P_{12}\Lambda_2Q_{21} \\ &\quad +(P_{11}\Lambda_1Q_{12}+P_{12}\Lambda_2Q_{22})P_{21}P_{11}^{-1} \\ &= A_{11}+A_{12}P_{21}P_{11}^{-1} \\ &\qquad (\because A=P\Lambda Q, \text{ which implies} \\ &\qquad A_{11}=P_{11}\Lambda_1Q_{11}+P_{12}\Lambda_2Q_{21} \text{ and} \\ &\qquad A_{12}=P_{11}\Lambda_1Q_{12}+P_{12}\Lambda_2Q_{22}) \text{ q.e.d.} \end{aligned}$$

This example might be useful when the target variable, which is included in the partitioned endogenous vector of y_t, y_{1t}, is a subset of the whole endogenous vector.

The above three examples do not have the explicit criterion function as a criterion of aggregation, but assume that the possession of a subset of characteristic values would be desirable for the approximation of dynamic behavior of the model.

Example 4 minimizes the aggregation error under the quadratic criteria.

(Example 4)[8]

Before discussing the aggregation, consider the following theorem by W. Fisher,

(Theorem 3) Given a matrix A of rank n, the matrix $\tilde{F}$ of rank l that minimizes

(3. 19) $$L=\text{tr}(\tilde{F}-A)'(\tilde{F}-A) \text{ is}$$

(3. 20) $\tilde{F}=R_1'R_1A$ where R_1 is the $l\times n$ matrix whose rows are the normalized characteristic vectors associated with the l largest characteristic roots of the matrix AA'.

(Proof) See Walter D. Fisher[8].

Note $\tilde{F}$ is the $n \times n$ matrix with rank l. Premultiplying R_1 in the equation (3.20),

(3.21) $$R_1\tilde{F}=R_1A$$

because the positive definite symmetric matrix, AA', has orthogonal characteristic vectors and can normalize as $R_1R_1'=I$.

Taking R_1 as an aggregation matrix,

(3.22) $$\underset{l\times l}{F}\ \underset{l\times n}{R_1}=R_1\tilde{F}=R_1A.$$

Postmultiplying R_1' in the equation (3.22)

(3.23) $$\underset{l\times l}{F}=R_1AR_1'.$$

Therefore, the aggregated model will be

(3.24) $$\underset{l\times 1}{Z_t}=\underset{l\times l}{F}\ \underset{l\times 1}{Z_{t-1}}+\underset{l\times m}{D}\ \underset{m\times 1}{x_t}$$

where $Z_t=R_1y_{t-1}$ and $D=R_1\ B$.

So far we have discussed aggregation by the characteristic vector, with several examples, which preserve the subset of the characteristic values. In reality, the characteristic vector is generally unknown. Therefore the above mentioned examples might be a criterion to keep in mind for the purpose of large-scale model control design, but may not be useful for the computational purposes in reality. More often than not, the aggregation matrix C, which is row-full rank, may be determined by the planner's a priori knowledge. The least square solution of the equation (2.4) can be considered, which implies a pseudo-inverse solution. In other words,

(3.25) $$F=CAC^{+}$$

where $$C^{+}=C'(CC')^{-1}$$

Clearly equation (3.25) is much more practical than the above examples where matrix C has been chosen to be orthogonal to some of characteristic vectors, and to get the form of equation (3.25) or (3.5).

When matrix C is determined a priori, we cannot expect that the rows of C and the characteristic vector of A are orthogonal, but equation(3.25) has its meaning in the sense of the least square solution. Thus the next section discusses aggregation under a certain class of a priori knowledge, or the nearly complete decomposability of the matrix.

4. Control on the Nearly Completely Decomposable System

At first, Ando-Simon's nearly completely decomposable system is reviewed briefly for its application to the control. Consider the autonomous system (for example, the Leontief matrix in the growth model),

(4.1)
$$\underset{n\times 1}{y_t} = \underset{n\times n}{A} \underset{n\times 1}{y_{t-1}}$$

The question on the aggregation of the model (4.1) will be, as in the last sections,

(4.2)
$$\underset{l\times 1}{Z_t} = \underset{l\times l}{F} \underset{l\times 1}{Z_{t-1}}$$

where
$$Z_t = \underset{l\times n}{C} y_t$$

Let us assume that the matrix may be represented by

(4.3)
$$A = A^* + \varepsilon A_*$$

where
$$A^* = \begin{bmatrix} A^*_1 & 0 & 0 \\ 0 & A^*_2 & 0 \\ 0 & 0 & A^*_l \end{bmatrix}$$

A^*_i is $n_i \times n_i$ matrix such that $\sum_{i=1}^{l} n_i = n$

ε is a very small real number, and A_* is an arbitrary matrix of the same dimension as A^*. Ando-Simon referred to matrices such as A as nearly decomposable matrices. Let us consider the following decomposition of A and A^*, if nonsingular,

(4.4)
$$A = P \Lambda P^{-1}$$

(4.5)
$$A^* = P^* \Lambda^* P^{*-1}$$

Furthermore, note that

(4. 5) $$\lim_{\varepsilon \to 0} P(\varepsilon) = P^*$$

(4. 7) $$\lim_{\varepsilon \to 0} \Lambda(\varepsilon) = \Lambda^*$$

$P(\varepsilon)$ and $\Lambda(\varepsilon)$ is a continuous function of ε and

(4. 8) $$\| P(\varepsilon) - P^* \| < \delta_1(\varepsilon)$$

and

(4. 9) $$\| \Lambda(\varepsilon) - \Lambda^* \| < \delta_2(\varepsilon)$$

where $\| \cdot \|$ denotes a norm of maximum absolute value. Under these circumstances and with a sufficiently small ε, Ando-Simon have shown that[1] in some finite time, T, or in the short-run, we may treat the system (4. 1) as though it consists of l independent subsystems and[2] in the long-run, we may look at our system as a set of relations among l aggregative variables, ignoring the relations within each of the subsystems. We have to note that in this case calculation of the aggregation matrix is much easier than in the examples of aggregation of Section 3. We don't have to compute the characteristic vectors of the whole matrix A but of the block diagonal matrix A^*, which can be computed by calculation of the characteristic vector of each subsystem A_i^*. Though this scheme is an approximate solution depending on ε, the aggregation matrix C would consist of orthogonal row vectors. Specifically, the aggregation matrix has the form,

(4. 10) $$\underset{l \times n}{C} = \begin{bmatrix} C_1 \cdots 0 \cdots 0 \\ \vdots \qquad\quad \vdots \\ 0 \cdots C_i \cdots 0 \\ \vdots \qquad\quad \vdots \\ 0 \cdots 0 \cdots C_l \end{bmatrix}$$

where C_i is an $1 \times n_i$ row vector.

Here another class of a priori information is to be considered. Suppose the planner could rearrange the matrix A by the row-column simultaneous permutation as

(4. 11) $$A = P \Lambda P^{-1}$$

where $$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}, \quad |\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|.$$

The matrix can be partitioned as

(4.12)
$$A=\left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right]\begin{array}{l} l \\ n-l \end{array}$$
$$\quad l \quad n-l$$

Suppose matrix A is nearly decomposable, in which A_{12} and A_{21} are sufficiently small. Furthermore, consider the similar transformation of matrix A as

(4.13)
$$A=P^{*-1}AP^{*}=\begin{bmatrix} \Lambda_1{}^{*} & P_1{}^{*-1}\ A_{12}\ P_2{}^{*} \\ P_2{}^{*-1}\ A_{21}\ P_1{}^{*} & \Lambda_2{}^{*} \end{bmatrix}$$
$$l \qquad n-l$$

where
$$A^{*}=\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}=\begin{bmatrix} P_1{}^{*}\ \Lambda_1{}^{*}\ P_1{}^{*-1} & 0 \\ 0 & P_2{}^{*}\ \Lambda_2{}^{*}\ P_2{}^{*-1} \end{bmatrix}$$
$$P^{*}=\begin{bmatrix} P_1{}^{*} & 0 \\ 0 & P_2{}^{*} \end{bmatrix}$$

and $\Lambda_1{}^{*}$ and $\Lambda_2{}^{*}$ are diagonal matrices whose diagonal elements are characteristic roots of A_{11} and A_{12}.

Milne[15] has defined the similarly transformed matrix $\tilde{A}$ as 'weakly coupled', if

(4.14)
$$\frac{|\lambda_{l+1}|}{|\lambda_l|}<1$$

(4.15)
$$\frac{kr\delta}{|\lambda_l|^2}<1$$

where k is min $(l,n-l)$, r is a maximum element of $P_1{}^{*-1}A_{12}P^{*}{}_2$ in absolute value and δ is a maximum element of $P_2{}^{*-1}A_{21}P_1{}^{*}$ in absolute value. Furthermore, the characteristic roots of A, Λ, can be approximated by

(4.16)
$$f_1(\lambda)=\det(\lambda I-\Lambda_1{}^{*})=0$$

(4.17)
$$f_2(\lambda)=\det(\lambda I-\Lambda_2{}^{*}+P_2{}^{*-1}A_{21}P_1{}^{*}\Lambda_1{}^{*}P_1{}^{*-1}A_{12}P_2{}^{*})=0$$

where det $(\cdot)$ denotes determinant.

In fact, conditions (4.14) and (4.15) can be checked by computation of the characteristic roots of the whole matrix A, which, on the contrary, is impractical for the purpose of analysis in this chapter. For this position, Milne has derived the necessary condition for conditions (4.14) and (4.15) as

(4.18)
$$\frac{r}{R}<1$$

where r denotes the minimum absolute root of the characteristic equation, $f_1(\lambda)$, in equation (4. 16) and R denotes the maximum absolute root of the characteristic equation, $f_2(\lambda)$, in equation (4. 17). In order to interpret this necessary condition for conditions (4. 14) and (4. 15) in the original matrix A, equations (4. 16) and (4. 17) can be rewritten as

(4. 19) $$f_1(\lambda)=\det(\lambda \boldsymbol{I}-\boldsymbol{A}_{11})=0$$

(4. 20) $$f_2(\lambda)=\det(\lambda \boldsymbol{I}-\boldsymbol{A}_{22}+\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})=0$$

because Λ_1^* and $(\Lambda_2^*-\boldsymbol{P}_2^{*-1}\boldsymbol{A}_{21}\boldsymbol{P}_1^*\Lambda_1^{*-1}\boldsymbol{P}_1^{*-1}\boldsymbol{A}_{12}\boldsymbol{P}_2^*)$ can be similarly transformed into $\boldsymbol{A}_{11}$ and $\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$. Therefore, equation (4. 1) can be separated into two different subsystems which have equations(4. 19) and(4. 20) as a characteristic equation,

(4. 21) $$\underset{l\times 1}{\boldsymbol{y}_{1t}}=\boldsymbol{A}_{11}\boldsymbol{y}_{1t-1}+\boldsymbol{V}_{1t}$$

(4. 22) $$\underset{(n-l)\times 1}{\boldsymbol{y}_{2t}}=(\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})\boldsymbol{y}_{2t-1}+\boldsymbol{V}_{2t}$$

where $$\boldsymbol{V}_{1t}=\boldsymbol{A}_{12}\boldsymbol{y}_{2t-1}$$

and $$\begin{aligned}\boldsymbol{V}_{2t}&=\boldsymbol{A}_{21}\boldsymbol{y}_{1t-1}+\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{y}_{2t-1}\\&=\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{y}_{1t}\end{aligned}$$

and then $\boldsymbol{V}_{1t}$ and $\boldsymbol{V}_{2t}$ can be treated as exogenous variables. Clearly the accuracy of two segregated subsystems, (4. 21) and (4. 22) depends upon the accuracy of the prediction of the exogenized variables V_{1t} and V_{2t}. Now let us consider a weakly coupled structural econometric model. The econometrician more frequently encountered with the structural model rather than the reduced form, a large-scale structural model being as well difficult to transform into the derived reduced form. Suppose the structural econometric model is given by

(4. 23) $$\boldsymbol{A}_0\boldsymbol{y}_t=\boldsymbol{A}_1\boldsymbol{y}_{t-1}+\boldsymbol{g}(t)$$

where $g(t)$ includes the exogenous variables, or in partitioned form,

(4. 23)′ $$\begin{matrix} l \\ n-l \end{matrix}\begin{bmatrix}\boldsymbol{A}_{0,11} & \boldsymbol{A}_{0,12}\\ \boldsymbol{A}_{0,21} & \boldsymbol{A}_{0,22}\end{bmatrix}\begin{bmatrix}\boldsymbol{y}_{1t}\\ \boldsymbol{y}_{2t}\end{bmatrix}=\begin{matrix} l \\ n-l \end{matrix}\begin{bmatrix}\boldsymbol{A}_{1,11} & \boldsymbol{A}_{1,12}\\ \boldsymbol{A}_{1,21} & \boldsymbol{A}_{1,22}\end{bmatrix}\begin{bmatrix}\boldsymbol{y}_{1t-1}\\ \boldsymbol{y}_{2t-1}\end{bmatrix}+\begin{bmatrix}\boldsymbol{g}_{1(t)}\\ \boldsymbol{g}_{2(t)}\end{bmatrix}$$
$$\quad\; l \quad n-l \qquad\qquad\qquad l \quad n-l$$

Let

(4.24)
$$A_0^{-1}=\begin{bmatrix} A_0^{11} & A_0^{12} \\ A_0^{21} & A_0^{22} \end{bmatrix}$$

and

(4.25)
$$A=A_0^{-1}A_1=\begin{bmatrix} A_0^{11}A_{1,11}+A_0^{12}A_{1,21} & A_0^{11}A_{1,12}+A_0^{12}A_{1,22} \\ A_0^{21}A_{1,11}+A_0^{22}A_{1,21} & A_0^{21}A_{1,12}+A_0^{22}A_{1,22} \end{bmatrix}$$

Analogously we could evaluate $f_1(\lambda)$ and $f_2(\lambda)$ of equation (4.19) and equation (4.20). Particularly we are interested in $f_1(\lambda)$ which would include the largest l characteristic roots approximately and $f_1(\lambda)$ has the rather simple form of

(4.26)
$$f_1(\lambda)=|\lambda I-(A_0^{11}A_{1,11}+A_0^{12}A_{1,21})|=0$$

with appropriate conditions like (4.14) and (4.15). From the partitioned matrix inverse, equaton (4.26) can be rewritten as

(4.27)
$$f_1(\lambda)=|\lambda(A_{0,11}-A_{0,12}A^{-1}{}_{0,22}A_{0,21}) -(A_{1,11}-A_{0,12}A^{-1}{}_{0,22}A_{1,21})|=0$$ [2)]

The model with the same characteristic roots as $f_1(\lambda)$will be

(4.28)
$$(A_{0,11}-A_{0,12}\,A^{-1}{}_{0,22}\,A_{0,21})\underset{l\times 1}{y_{1t}} =(A_{1,11}-A_{0,12}A^{-1}{}_{0,22}A_{1,21})\underset{l\times 1}{y_{1t-1}} +V_{1t}$$

where V_{1t} is a function of y_{2t-1} and $g(t)$ and the coefficient of y_{2t-1} in V_{it} is $(A_{1,12}-A_{0,12}A^{-1}{}_{0,22}A_{1,22})$. From equation (4.28), we could infer that a structural econometric model with a distributed lag structure can be simplified in an analogous way. Suppose the distributed lag structural econometric model is given by

(4.29)
$$A(L)y_t=g(t)$$

where
$$A(L)=A_0+A_1L+\cdots+A_KL^K$$

and L denotes a lag operator.

Let $A(L)$ be partitioned by

2) $f_2(\lambda)$ can be constructed in a similar way to the equation (4.20) and the necessary condition for the weak coupling, (4.18), will also hold with respect to $f_1(\lambda)$ and $f_2(\lambda)$.

(4.30) $$A(L)=\begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix}$$

Then the simplified model will be

(4.31) $$\{A_{11}(L)-A^{-1}_{0,12}A_{0,22}A_{21}(L)\}\,y_{1t}=V_{it}$$

So far we have discussed the simplification of the model by Ando-Simon's nearly completely decomposable system and Milne's weakly coupled system. The former requires a priori knowledge of the sufficiently smallness condition of the interaction term between subsystems of the off-diagonal blocks in the coefficient matrix of the lagged endogenous variables of the reduced form. The latter requires a subsystem to dominate other subsystem in the sense of characteristic roots in addition to the sufficiently smallness condition on the interaction term. Though the distributed lag model might be analyzed in a similar way and the dominating subsystem can be constructed by equation (4.31), it is very hard to determine the conditions for weak coupling or nearly complete decomposability. This is so because the off-diagonal blocks (interactions) are described by the polynomial with respect to the lag operator L rather than by constant coefficients in case of the first-order system. A similar problem arises when the formulation of equation (4.1) includes the instrument variables, for control of the econometric model can be interpreted as a procedure of endogenization of controllable exogeous variables by the same criteria. This is described as a loss function or a welfare function and as such the interpretation implies that the coefficient matrix of the instruments should receive the foregoing analysis, even though in somewhat different form. As a counterpart of equation (4.1) we can rewrite the equation (1.2) as

(4.1)′ $\boldsymbol{y}_t=\boldsymbol{A}\boldsymbol{y}_{t-1}+\boldsymbol{B}\boldsymbol{x}_t$

or $$\begin{bmatrix} \boldsymbol{y}_{1t} \\ \boldsymbol{y}_{2t} \end{bmatrix}=\begin{matrix} l \\ n-l \end{matrix}\underset{\quad l \quad n-l}{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}\begin{bmatrix} \boldsymbol{y}_{1t-1} \\ \boldsymbol{y}_{2t-1} \end{bmatrix}+\begin{matrix} l \\ n-l \end{matrix}\underset{m}{\begin{bmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \end{bmatrix}}\boldsymbol{x}_t$$

The solution form will be

$$(4.1)'' \quad \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{t} \begin{bmatrix} y_{1,0} \\ y_{2,0} \end{bmatrix} + \sum_{\tau=1}^{t} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{t-\tau} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} x_\tau$$

So far we have discussed about first part (autonomous part) of the decomposition of the system. This argument may not hold for the system of (4.1)'' because the second part includes matrix B by multiplication.

For this purpose, let us consider the second form of the canonical representation (1.11) and then the solution for be

$$(4.35) \quad \begin{bmatrix} Z_{1t} \\ Z_{2t} \end{bmatrix} = \underset{l \quad n-l}{\begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}}^{t} \begin{bmatrix} Z_{10} \\ Z_{20} \end{bmatrix} + \sum_{\tau=1}^{t} \underset{l \quad n-l}{\begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}}^{t-\tau} \begin{pmatrix} I \\ 0 \\ \cdots \\ 0 \end{pmatrix} x_\tau$$

where $Z_t = S^{-1} y_t$, $\tilde{A} = S^{-1}AS$, and $\underset{n\times n}{S^{-1}} \underset{n\times m}{B} = \binom{I}{0}$, assuming l is greater than m without loss of generality.

With the canonical representation of equation (4.35), sufficient smallness of the off-diagonal blocks, $\tilde{A}_{12}$ and $\tilde{A}_{21}$, of the decomposition satisfies the first part and the second part at the same time. The foregoing analysis also depends on the choice of matrix S. Indeed, it is not easy to choose the appropriate S^{-1} to get the off-diagonal block sufficiently small and to be orthogonal to the column vectors of matrix B. Furthermore, the choices of matrix S^{-1} orthogonal to the column vectors of matrix B is infinite because S can be constructed by arbitrary $(n-m)$ linearly independent column vectors and m linearly independent column vectors of matrix B. If matrix S is constructed by

$$(4.36) \quad S = \begin{matrix} m \\ n-m \end{matrix} \overset{m}{\underset{m \quad n-m}{\begin{bmatrix} B_1 & 0 \\ B_2 & I \end{bmatrix}}}$$

and the similar transformation of matrix A by such S satisfy the sufficient smallness of the off-diagonal blocks in the equation (4.35), the construction of equation (4.35) is not too difficult because S^{-1} can be computed simply by

(4.37)
$$S^{-1}=\begin{array}{c} m \\ n-m \end{array}\begin{bmatrix} B_1^{-1} & 0 \\ -B_2B_1^{-1} & I \end{bmatrix}$$
$$\qquad\qquad m \qquad n-m$$

which requires the inversion of $m\times m$ matrix expected to be fairly small enough. At any rate there is no clear-cut method for the choice of matrix S, while the S of (4.36) and (4.37) is simply for computational consideration. When the matrices A and B are polynomials with respect to the lag operator, analysis is much more difficult. Thus, for practical application, multiplier calculation would give much better insight for decomposition of the system. In fact, this might be the only way for a large-scale nonlinear model. To see this point, let us consider the most general model as

(4.38)
$$\underset{n\times 1}{y_t} = \underset{n\times n}{A(L)}\ \underset{n\times 1}{y_{t-1}}+\underset{n\times m}{B(L)}\ \underset{m\times 1}{x_t}$$

where $A(L)=A_1+A_2L+\cdots+A_pL^{p-1}$

and $B(L)=B_0+B_1L+\cdots+B_qL^q$

Then explicit expressions for the matrices of impact and interim multipliers are given by[3]

(4.39) for impact multipliers: $\underset{n\times m}{B_0}$

for first period multipliers: $B_1+M(1)B_0$

for second period multipliers: $B_2+M(1)B_1+M(2)B_0$

for qth period multipliers: $\sum_{\tau=0}^{q} M(\tau)B_{q-\tau}$

for $q+1$th period multipliers: $\sum_{\tau=0}^{q} M(\tau+1)B_{q-\tau}$

where

(4.40)
$$\underset{n\times m}{M(t)}=A_1M(t-1)+A_2M(t-2)+\cdots+A_pM(t-p)$$
$$=\sum_{\tau=1}^{p} A_\tau M(t-\tau)$$

$M(t)=I$ when $t=0$

and $M(t)=0$ when $t<0$

We have to note that the matrix $M(t)$ determines internal dynamics and expression (4.39) determines the effect of the endogenous variables by instrument dynamics which is a composite function of internal and external dynamics. From (4. 39) and (4. 40) it is clear that multiplier calculation of $p+q$ periods would at least give good insight into the probability of decomposition of the model, without the canonical representation of the second form(1. 11).

REFERENCES

[1] Ando, Albert and Franklin M. Fisher, "Near-decomposability, partition and aggregation and the relevance of stability discussion," *International Economic Beview*, Jan. 1963.

[2] ——, and Simon, Herbert A. "Aggregation of variables in dynamic systems," *Econometrica*, April 1961.

[3] Brissimis, S.N., "Multiplier effects for higher than first order linear dynamic econometric models," *Econometrica*, May 1976.

[4] Davison, E.J., "Method for simplifying linear dynamic systems," *IEEE Transactions on Automatic Control*, January 1966.

[5] Engle, Robert F. and Ta-Chung Liu, "Effects of aggregation over time on dynamic characteristics of an econometric model," *Econometric Models of Cyclical Behavior*, edited by Bert G. Hickman, Studies in Income and Wealth, No. 36, NBER.

[6] Fair, R.C., "On the solution of optimal control problems as maximization problems," *Annals of Economic and Social Measurement*, Vol. 3, No. 1, Jan. 1974.

[7] Fisher, F.M., "On the cost of approximate specification in simultaneous equation estimation," *Econometrica*, April 1961.

[8] Fisher, W.D., "Optimal aggregation in multi-equation prediction models," *Econometrica*, October, 1962. (a)

——, "Estimation in the linear decision model," *International Economic Review*, January 1962. (b)

[9] Gantmacher, F.R., *Matrix Theory,* Chelsea, 1960, 215-220.

[10] Holbrook, R.S., "A practical method for controlling a large nonlinear stochastic system," *Annals of Economic and Social Measurement,* 1974, Vol. 3, 155-176.

[11] Lange, Oscar, *Price Flexibility and Employment,* 1952.

[12] Luenberger, D.G., "Canonical forms for linear multivariate systems," *IEEE Transactions on Automatic Control,* June 1967.

[13] McKenzie, L., "Matrices with dominant diagonals and economic theory," in Arrow, et al., *Mathematical Methods in Social Sciences,* 1959, 47-62.

[14] Marcus, M. and M. Henryk, *A Survey of Matrix Theory and Matrix Inequalities,* Boston, 1964.

[15] Milne, R.D., "The analysis of weakly coupled dynamical systems," *International Journal of Control,* August 1965.

[16] "Special issue on control theory applications," *Annals of Economic and Social Measurement,* Oct. 1972, Jan. 1974, Spring 1976, Summer 1976.

[17] Theil, H., "The aggregation implications of identifiable structural macrorelations," *Econometrica,* Jan. 1959.

[18] Vargar, R.S., "On recurring theorems on diagonal dominance," *Linear Algebra and its Applications,* 1976, 1-9.