ANDREI ROGERS

On Perfect Aggregation in the Matrix Cohort Survival Model of Interregional Population Growth

Institute of Urban & Regional Development
University of California, Berkeley

Reprinted from Journal of Regional Science
Vol. 9, No. 3, 1969

JOURNAL OF REGIONAL SCIENCE, VOL. 9, NO. 3, 1969

ON PERFECT AGGREGATION IN THE MATRIX COHORT-SURVIVAL MODEL OF INTERREGIONAL POPULATION GROWTH*

RP 58

Andrei Rogers†

1. INTRODUCTION

Users of multisectoral growth models have long recognized the importance of the aggregation problem and the fact that the results of the analysis depend on the particular consolidation scheme which is used to combine industries, regions, and time. Several authors have considered the theoretical and practical aspects of the aggregation problem in input-output analysis.[1] However, no comparable literature exists for the corresponding problem in demographic analysis. The purpose of this note is to show that conditions for "perfect aggregation," analogous to those developed by economists for the matrix input-output model, also can be derived for the matrix cohort-survival model.[2]

2. AGGREGATION IN THE INTERREGIONAL MATRIX MODEL OF POPULATION GROWTH AND DISTRIBUTION

2.1 *Consolidation and Deconsolidation Operations in Matrix Form*

Consolidation and deconsolidation operations may be conveniently carried out by simple matrix multiplication. For example, to consolidate a three-dimensional vector, v, into the two-dimensional vector, $\hat{v}$, such that the first element of $\hat{v}$ is the sum of the first two elements of v, we define the *consolidation operator:*

$$C = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

such that

$$\hat{v} = Cv = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 + v_2 \\ v_3 \end{bmatrix}.$$

---

* This research has been supported in part by the U. S. Economic Development Administration.

† The author is Associate Professor of City and Regional Planning at the University of California, Berkeley.

1 See, for example, items [1] through [10] listed in the References.

2 For a description of the interregional matrix cohort-survival model, see Rogers [11].

Date received: October, 1968; revised, May, 1969.

To reverse this consolidation, we may define the *deconsolidation operator:*

$$D = \begin{bmatrix} d_{11} & d_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where d_{11} and d_{12} are a system of weights such that $d_{11} = v_1/(v_1 + v_2)$ and $d_{12} = 1 - d_{11}$. Thus,

$$v = D'\hat{v} = \begin{bmatrix} \dfrac{v_1}{v_1 + v_2} & 0 \\ \dfrac{v_2}{v_1 + v_2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 + v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}.$$

Note that $CD' = I$.

We may generalize the above results by defining the consolidation operator as:

$$(1) \qquad \underset{(m\times n)}{C} = \begin{bmatrix} c'(1) & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & c'(2) & 0 & & & & & & \cdot \\ \cdot & \cdot & \cdot & & & & & & \cdot \\ \cdot & \cdot & \cdot & & & & & & \cdot \\ \cdot & \cdot & \cdot & & & & & & \cdot \\ \cdot & \cdot & \cdot & & & & & & 0 \\ 0 & 0 & \cdot & & & & & & c'(m) \end{bmatrix}$$

where $c'(i)$ is the $s(i)$-dimensional vector,

$$c'(i) = [1 \quad 1 \cdots 1]; \qquad (i = 1, 2, \cdots, m).$$

Thus if

$$\underset{(n\times 1)}{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

then, for an appropriately defined C,

$$(2) \qquad \underset{(m\times 1)}{\hat{v}} = \begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \\ \vdots \\ \hat{v}_m \end{bmatrix} = \underset{(m\times n)}{C} \ \underset{(n\times 1)}{v}.$$

The corresponding deconsolidation operator is defined as

$$(3) \qquad \underset{(m\times n)}{D} = \begin{bmatrix} d'(1) & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & d'(2) & 0 & & & & & & \cdot \\ \cdot & \cdot & & & & & & & \cdot \\ \cdot & \cdot & & & & & & & \cdot \\ \cdot & \cdot & & & & & & & \cdot \\ \cdot & \cdot & & & & & & & 0 \\ 0 & 0 & & & & & & & d'(m) \end{bmatrix}$$

where $d'(i)$ is the $s(i)$-dimensional vector of weights,

$$d'(i) = [d_{i1}d_{i2} \cdots d_{is(i)}], \qquad (i = 1, 2, \cdots, m)$$

and

$$d_{ij} = \frac{v_i}{\sum_{j=k}^{k+s(i)} v_j},$$

the denominator being the sum of the $s(i)$ elements of v that are to form $\hat{v}_i$. It follows then that, for an appropriately defined D,

(4) $$\underset{(n\times 1)}{v} = \underset{(n\times m)}{D'} \underset{(m\times 1)}{\hat{v}} .$$

Observe, once again, that

(5) $$\underset{(m\times n)}{C} \underset{(n\times m)}{D'} = \underset{(m\times m)}{I} .$$

2.2 *Perfect Aggregation in Multisectoral Growth Models*

If our fundamental *unconsolidated* multisectoral growth model is of the form

(6) $$v(t+1) = Gv(t),$$

and the corresponding *consolidated* growth model is

(7) $$\hat{v}(t+1) = \hat{G}\hat{v}(t),$$

then $\hat{G}$ may be obtained as a function of G and the *consolidation* and *deconsolidation* operators C and D. First, we recall (2) and (4) to establish that

(8) $$\hat{v}(t) = Cv(t),$$

and

(9) $$v(t) = D'(t)\hat{v}(t),$$

where D now receives the temporal subscript of the vector to which it is applied, thereby reflecting the point in time when the proportional weights are computed. Next, we observe that

(10) $$\begin{aligned} \hat{v}(t+1) &= Cv(t+1), \text{ [by (8)]}, \\ &= CGv(t), \text{ [by (6)]}, \\ &= CGD'(t)\hat{v}(t), \text{ [by (9)]}, \\ &= \hat{G}\hat{v}(t), \end{aligned}$$

where

(11) $$\hat{G} = \underset{(m\times n)}{C} \underset{(n\times n)}{G} \underset{(n\times m)}{D'} (t).$$

We may now identify those consolidation schemes which will produce *perfect aggregation*, that is, a consolidated model which produces results that would have been obtained in the absence of any consolidation. First, we note that both (10) and (6) yield "errorless" results for $\hat{v}(t+1)$ and $v(t+1)$, respectively, since they are the products of the unconsolidated model. Hence

$$\hat{v}(t+1) = Cv(t+1) = CGv(t), \text{ [by (6)]} \tag{12}$$

and

$$v(t+1) = Gv(t) = GD'(t)\hat{v}(t), \text{ [by (9)]} \tag{13}$$

may be defined as the errorless consolidated and unconsolidated values, respectively, of the vector v at time $t+1$. Corresponding to each of these errorless results, we also have the potentially erroneous output of the consolidated model of (7). Thus, associated with errorless results of (12) and (13), respectively, are the potentially erroneous projections:

$$\hat{v}(t+1) = \hat{G}\hat{v}(t) = \hat{G}Cv(t), \text{ [by (8)]} \tag{14}$$

and

$$\begin{aligned} v(t+1) &= D'(t+1)\hat{v}(t+1), \text{ [by (4)]} \\ &= D'(t+1)\hat{G}\hat{v}(t), \text{ [by (7)].} \end{aligned} \tag{15}$$

Denoting the difference between the errorless and potentially erroneous projections by the vector ϕ, we use (12) and (14) to define

$$\begin{aligned} \phi_1 &= CGv(t) - \hat{G}Cv(t) \\ &= [CG - \hat{G}C]v(t), \end{aligned} \tag{16}$$

and (13) and (15) to define

$$\begin{aligned} \phi_2 &= GD'(t)\hat{v}(t) - D'(t+1)\hat{G}\hat{v}(t) \\ &= [GD'(t) - D'(t+1)\hat{G}]\hat{v}(t). \end{aligned} \tag{17}$$

It follows, therefore, that *perfect aggregation* occurs when ϕ_1 or ϕ_2 are null vectors, and this will happen if

$$CG = \hat{G}C \text{ [Consolidation Rule I]} \tag{18}$$

or

$$GD'(t) = D'(t+1)\hat{G} \text{ [Consolidation Rule II]} \tag{19}$$

We shall refer to (18) and (19) as the ***sufficient conditions for perfect aggregation in multisectoral growth models***. To illustrate their application, let us turn to a three-region population model and derive the conditions for perfect aggregation.

2.3 *Perfect Aggregation in the Matrix Cohort-Survival Model*

Imagine a three-region demographic system and assume that each region's population may be denoted by a 4-dimensional vector. Thus we have the three 4-dimensional vectors: w_i, w_j, and w_k, and may denote their change over time by:

$$w(t+1) = Gw(t), \tag{20}$$

where

$$G = \left[\begin{array}{cccc|cccc|cccc}
0 & b_{i2} & b_{i3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
s_{i1} & 0 & 0 & 0 & {}_1g_{ji} & 0 & 0 & 0 & {}_1g_{ki} & 0 & 0 & 0 \\
0 & s_{i2} & 0 & 0 & 0 & {}_2g_{ji} & 0 & 0 & 0 & {}_2g_{ki} & 0 & 0 \\
0 & 0 & s_{i3} & 0 & 0 & 0 & {}_3g_{ji} & 0 & 0 & 0 & {}_3g_{ki} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & b_{j2} & b_{j3} & 0 & 0 & 0 & 0 & 0 \\
{}_1g_{ij} & 0 & 0 & 0 & s_{j1} & 0 & 0 & 0 & {}_1g_{kj} & 0 & 0 & 0 \\
0 & {}_2g_{ij} & 0 & 0 & 0 & s_{j2} & 0 & 0 & 0 & {}_2g_{kj} & 0 & 0 \\
0 & 0 & {}_3g_{ij} & 0 & 0 & 0 & s_{j3} & 0 & 0 & 0 & {}_3g_{kj} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{k2} & b_{k3} & 0 \\
{}_1g_{ik} & 0 & 0 & 0 & {}_1g_{jk} & 0 & 0 & 0 & s_{k1} & 0 & 0 & 0 \\
0 & {}_2g_{ik} & 0 & 0 & 0 & {}_2g_{jk} & 0 & 0 & 0 & s_{k2} & 0 & 0 \\
0 & 0 & {}_3g_{ik} & 0 & 0 & 0 & {}_3g_{jk} & 0 & 0 & 0 & s_{k3} & 0
\end{array}\right]$$

$$= \left[\begin{array}{c|c|c}
G_{ii} & G_{ji} & G_{ki} \\ \hline
G_{ij} & G_{jj} & G_{kj} \\ \hline
G_{ik} & G_{jk} & G_{kk}
\end{array}\right],$$

$$w(t) = \left[\begin{array}{c} w_i(t) \\ \hline w_j(t) \\ \hline w_k(t) \end{array}\right],$$

and

$w_{ir}(t)$ = the population in the r^{th} age group, in the i^{th} region, at time t;

b_{ir} = the number of surviving births, per person, attributable to the r^{th} age group, in the i^{th} region, during the time period $(t, t + 1)$; and

s_{ir} = the survivorship proportion of the r^{th} age group, in the i^{th} region, during the time period $(t, t + 1)$.

${}_rg_{ij}(t)$ = total number of migrants in the r^{th} age group, in the i^{th} region, who move to the $(r + 1)^{\text{st}}$ age group in the j^{th} region during the time interval $(t, t + 1)$.

Now consider the consequences of consolidating regions i and j. First, we define the consolidation operator, C:

$$C = \left[\begin{array}{c|c|c} I & I & 0 \\ \hline 0 & 0 & I \end{array}\right] \tag{21}$$

and the corresponding deconsolidation operator, $D(t)$:

$$D(t) = \left[\begin{array}{c|c|c} D^*(t) & I - D^*(t) & 0 \\ \hline 0 & 0 & I \end{array}\right], \tag{22}$$

where $D^*(t)$ is a diagonal submatrix with elements

$$d_h = \frac{w_{ih}(t)}{w_{ih}(t) + w_{jh}(t)}.$$

Given the growth matrix G and the operators C and D, we may use (14) to derive the consolidated growth matrix:

$$\hat{G} = CGD'(t)$$

(23)
$$= \begin{bmatrix} I & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} G_{ii} & G_{ji} & G_{ki} \\ G_{ij} & G_{jj} & G_{kj} \\ G_{ik} & G_{jk} & G_{kk} \end{bmatrix} \begin{bmatrix} D^*(t) & 0 \\ I - D^*(t) & 0 \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} (G_{ii} + G_{ij})D^*(t) + (G_{jj} + G_{ji})[I - D^*(t)] & G_{ki} + G_{kj} \\ G_{ik}D^*(t) + [I - D^*(t)]G_{jk} & G_{kk} \end{bmatrix}.$$

With these redefinitions of G, C, $D(t)$, and $\hat{G}$, we may invoke Consolidation Rules I and II, as defined by Equations (18) and (19), to express the sufficient conditions for perfect aggregation to be possible.

2.4 *The Joint Impact of Several Consolidations*

Having considered the consequences of spatial aggregation, let us now extend our results to include the effects of multidimensional aggregation by applying the consolidation and deconsolidation operators for each form of aggregation in sequence. An example will clarify the method.

Imagine a three-region, two-color, and four-age-group demographic system. Such a population may be denoted by a column vector containing 24 elements, and its growth regime may be defined by a 24-by-24 matrix. The fundamental matrix growth equation for this population system, as before, is:

(24)
$$w(t + 1) = Gw(t),$$

where the matrix G now consists of submatrices arranged as follows:

$$G = \begin{bmatrix} G^m_{ii} & G^m_{ji} & G^m_{ki} & 0 & 0 & 0 \\ G^m_{ij} & G^m_{jj} & G^m_{kj} & 0 & 0 & 0 \\ G^m_{ik} & G^m_{jk} & G^m_{kk} & 0 & 0 & 0 \\ 0 & 0 & 0 & G^n_{ii} & G^n_{ji} & G^n_{ki} \\ 0 & 0 & 0 & G^n_{ij} & G^n_{jj} & G^n_{kj} \\ 0 & 0 & 0 & G^n_{ik} & G^n_{jk} & G^n_{kk} \end{bmatrix}.$$

The superscripts m and n differentiate white and nonwhite growth operators, respectively, and the subscripts i, j, and k differentiate regional growth operators.

Now let us consolidate our demographic system into one consisting of two regions, $(i + j)$ and k; a single total population; and a doubled time interval. We begin by performing the temporal consolidation first, by replacing the growth operator with its square. The order of each successive consolidated growth matrix is listed below it, and each consolidation and deconsolidation operator is distinguished by a subscript. Thus

$$\underset{(12\times 12)}{\hat{G}} = C_t G^2 D_t'(t).$$

Next we aggregate the white and nonwhite portions of the population:

$$\underset{(6\times 6)}{\hat{\hat{G}}} = C_c \hat{G} D_c'(t) = C_c C_t G^2 D_t'(t) D_c'(t).$$

Finally, we consolidate regions i and j:

(25)
$$\underset{(4\times 4)}{\hat{\hat{\hat{G}}}} = C_r \hat{\hat{G}} D_r'(t) = C_r C_c C_t G^2 D_t'(t) D_c'(t) D_r'(t)$$
$$= CG^2 D'(t),$$

where

$$C = C_r C_c C_t \quad \text{and} \quad D'(t) = D_t'(t) D_c'(t) D_r'(t).$$

The consolidation rules for this particular multidimensional aggregation, therefore, are:

(26) $$CG^2 = \hat{\hat{\hat{G}}}C \text{ [Consolidation Rule I]}$$

and

(27) $$G^2 D'(t) = D'(t+2)\hat{\hat{\hat{G}}} \text{ [Consolidation Rule II]}$$

As before, we distinguish between the deconsolidation operator D at the beginning and end of the time interval.

Thus, we have identified the two rules for consolidation which, if strictly met, will produce *perfect aggregation* in the matrix cohort-survival model of interregional population growth. The first states that two regions or populations whose age-specific rates of fertility and mortality in the absence of intermigration and whose age-specific rates of outmigration to the unconsolidated region or regions are equal can be perfectly aggregated. The second asserts that two regions or populations with different age-specific rates can be perfectly aggregated if their population distributions retain a constant proportional relationship to each other. This latter consolidation rule is likely to be of interest only in stable population theory.

3. CONCLUSION

In this paper we have posed the question: When is it possible to consolidate regions, people, and time in a matrix cohort-survival model of population growth and still obtain the same results as would have been obtained by using the original unconsolidated model? Our answer to this question has been "almost never." The special conditions when such perfect aggregation is possible are not likely to be met in practice; therefore, it becomes necessary to consider what might be called "good" aggregation procedures, that is, procedures which reflect the objective of "best possible" accuracy, rather than "perfect" accuracy.

Although the aggregation problem has received considerable attention from scholars of input-output models, relatively little work has been done toward suggesting criteria for consolidation procedures. The need for developing such criteria has been recognized, but only a few specific suggestions have been made. Fisher [6] has proposed the use of a *minimal distance criterion* that seems to provide a good index with which to differentiate and evaluate alternative consolidation schemes. Charnes and Cooper [4] suggest the use of the *norm* of the matrix $CG - \hat{G}C$ as an appropriate measure. Finally, Ijiri [9] develops the concept of a *linear aggregation coefficient* and illustrates how it might be used as a measure of the accuracy of alternative consolidation schemes.

The *minimal distance criterion*, the *norm* of $CG - \hat{G}C$, and the *linear aggregation coefficient* all can be applied to cohort-survival models of population growth to indicate how closely a particular consolidation scheme meets the conditions of perfect aggregation. Their effectiveness in such situations, however, is yet to be tested with empirical data.

REFERENCES

[1] Ara, K. "The Aggregation Problem in Input-Output Analysis," *Econometrica*, 26 (1959), 257–262.

[2] Balderston, J. B. and T. M. Whitin. "Aggregation in the Input-Output Model," in O. Morgenstern, ed., *Economic Activity Analysis*. New York: John Wiley and Sons, Inc., 1954, pp. 79–128.

[3] Barna, T. "Classification and Aggregation in Input-Output Analysis," in T. Barna, ed., *The Structural Interdependence of the Economy*. Proceedings of an International Conference on Input-Output Analysis, Varenna, Italy, 1954. New York: John Wiley and Sons, Inc., 1954, pp. 175–185.

[4] Charnes, A. and W. W. Cooper. *Management Models and Industrial Applications of Linear Programming*, I. New York: John Wiley and Sons, Inc., 1961, pp. 448–467.

[5] Fei, J. C. H. "A Fundamental Theorem for the Aggregation Problem of Input-Output Analysis," *Econometrica*, 24 (1956), 400–412.

[6] Fisher, W. D. "Criteria for Aggregation in Input-Output Analysis," *Review of Economics and Statistics*, 40 (1958), 250–260.

[7] Hatanaka, M. "Note on Consolidation Within a Leontief System," *Econometrica*, 20 (1952), 301–303.

[8] Holzman, M. "Problems of Classification and Aggregation," in W. Leontief, ed., *Studies in the Structure of the American Economy*. New York: Oxford University Press, 1953, Chap. 9.

[9] Ijiri, Y. "The Linear Aggregation Coefficient as the Dual of the Linear Correlation Coefficient," *Econometrica*, 36 (1968), 252–259.

[10] Malinvaud, E. "Aggregation Problems in Input-Output Models," in T. Barna, ed., *The Structural Interdependence of the Economy*. Proceedings of an International Conference on Input-Output Analysis, Varenna, Italy, 1954. New York: John Wiley and Sons, Inc., 1954, pp. 189–202.

[11] Rogers, A. *Matrix Analysis of Interregional Population Growth and Distribution*. Berkeley: University of California Press, 1968.