MATRIX MULTIPLICATION AS A TECHNIQUE OF POPULATION ANALYSIS

NATHAN KEYFITZ

This paper takes advantage of the fact that the changes in numbers of a human (or any other) population may be traced by simple matrix multiplication. When an appropriately constructed matrix M is repeatedly multiplied by an age distribution it carries the numbers at the several ages into successive time periods. The age distribution n periods after the starting point is M^n multiplied by the initial age distribution. This property of the matrix has several potential uses in demography; in the present paper we confine ourselves to a few considerations regarding individual elements and columns of M^n.

While a matrix is an array of numbers, it can be treated in a manner analogous to a single number and many of the propositions of ordinary algebra apply to it. For instance, if S has the same number of rows and columns as B, then it is permissible to construct $M = S + B$ by simply adding the elements in corresponding positions. For multiplication the convention is to take rows by columns. Thus to calculate $A \times B$ we need merely know that the jth element of its ith row is constructed by multiplying the several elements of the ith row of A by the corresponding elements of the jth column of B, and adding these products.

The usefulness of population projections as a means of demographic study does not depend on their constituting a good prediction of what will actually happen to numbers in the several age groups. This is evidently true when the projection is made on the basis of the birth and death rates of a given year, but it is true even when the rates prevailing in the past are assumed to change in the future. The point is discussed in definitive fashion by Professor Donald J. Bogue in "The Population of the United States."[1] In the present article a most abstemious viewpoint is taken in relation to prediction. The projection is considered merely as a way of regarding the fertility and mortality of a single past year. We do not think of it as estimating the population that would exist if the rates of that year were to continue into the future any more than this is brought to mind when any other rate from last year's statistics is calculated; few calculators of such rates would consider it necessary to disavow explicitly the assumption that the rates would continue into the future. The projection procedure is here considered with as little attention to prediction as the simple calculation of the increase of the population from last year to this.

THE PROBLEM: TO SUMMARIZE THE FERTILITY AND SURVIVORSHIP PATTERNS OF A POPULATION

The projection procedure is separable from the age distribution of the original population, and is considered as an operator whose elements reflect the regime of mortality and fertility of the given population in the given year, and nothing more. This is not a novel approach, for it appears in at least three papers prior to this one. Bernardelli used it to investigate what might happen in a hypothetical population of beetles.[2] P. H. Leslie worked it out for a real population of the females of a domesticated brown rat stock housed at the Wistar Institute, Philadelphia, for which a life table was available.[3] Alvaro Lopez did not apply it to actual data, but considered with its help some important problems in demographic theory.[4] The matrix, even though it has several elements, is considered in what follows as a number, and one which has a unique claim to represent the essential facts of life and death in any human population. It is a multidimensional analogue of the ratio of this year's population to last year's, differing only from this single figure in that it takes account of age and separates the effect of births from that of deaths.

The search for some simple way of summarizing the pattern of fertility and mortality to which a population is subject has been a constant preoccupation of demographers. Crude birth and death rates, standardized rates, life tables, gross and net reproduction rates, the intrinsic rates of natural increase, have been developed, applied, and their limitations pointed out. The matrix of which this paper makes use is a special combination of fertility and survivorship containing 15 non-zero elements, which is a disadvantage in that it is more cumbersome to handle 15 numbers than a single number. Our matrix is not as readily manipulated on a desk calculator as a simple net reproduction rate but, on the other hand, it gives the meaning of past rates for the trajectory of a population not only after stability has been attained, but also on the path to stability. The Net Reproduction Rate (or its refinement, the intrinsic rate of natural increase) tells us what will be the consequences for population growth if the existing rates of survivorship and maternity are continued indefinitely; the matrix enables us to deal with the practically more important question of the consequences of the present rates for the immediate future as well. In order to keep the tables here presented within reasonable length, we use five-year age groups and confine our attention to that group to which the Net Reproduction Rate refers—females under 45 years of age—the most interesting part of a population from the viewpoint of reproduction.

THE CONVENTIONAL POPULATION PROJECTION

The form of the matrix is suggested by the ordinary population projection, in which explicit assumptions are made as to future births and deaths. Whether such a projection may be best looked on as a prediction or as a mere exercise in working out the consequences of the given assumptions depends on the taste of those who are making the calculations and using them. In the present instance we choose to regard the matrix and its products as a form of analysis of the specific data referring to the immediate past from which they are constructed. The future will be referred to only as a manner of describing and summarizing data for the past.

To make the discussion concrete, consider the 1960 population of the United States as counted in the census of April 1. Excluding armed services overseas and civilians absent for extended periods of time makes the total count 179,323,175. Our first simplification is to deal only in thousands; our second, as was noted, is to confine the study to women under 45 years of age. This gives as the jumping-off point the age distribution shown in column 2 of Table 1.

TABLE 1. UNITED STATES FEMALE POPULATION, UNDER 45 YEARS OF AGE, ROUNDED TO THOUSANDS, AS COUNTED IN THE CENSUS OF APRIL 1, 1960, AND PROJECTED BY THE COMPONENTS METHOD TO APRIL 1, 1965.

(1)	*1960* (2)	*1965* (3)	*Average 1960–1965* (4)	*Age-Specific Birth Rate 1960* (5)	(4)×(5) (6)
Total	63,506,000	67,853,000			
0– 4	9,991,000	10,531,000			
5– 9	9,187,000	9,954,000			
10–14	8,249,000	9,171,000			
15–19	6,586,000	8,232,000	7,409,000	.0899	666,000
20–24	5,528,000	6,566,000	6,047,000	.2581	1,561,000
25–29	5,536,000	5,506,000	5,521,000	.1974	1,090,000
30–34	6,103,000	5,507,000	5,805,000	.1127	654,000
35–39	6,402,000	6,057,000	6,229,000	.0562	350,000
40–44	5,924,000	6,329,000	6,126,000	.0164	100,000
					4,421,000

I have placed in the same table the projection by the ordinary components method. The United States life table for 1960 was used to secure the probabilities of surviving from each age group to the following age group during the five-year period; for the girl children 0–4 in 1960 this probability was .99633; for those 5–9 it was .99829, etc. A matrix operator which would premultiply the 1960 age distribution to put survivors into the next age groups is as follows:

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .99633 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .99829 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .99789 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .99689 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .99606 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .99477 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .99253 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .98867 & 0 \end{bmatrix}$$

where the only non-zero terms are in the subdiagonal, and these are the successive values of the life-table function $\frac{{}_5L_{x+5}}{{}_5L_x}$, i.e., $\frac{{}_5L_5}{{}_5L_0}$, $\frac{{}_5L_{10}}{{}_5L_5}$, etc., ending with $\frac{{}_5L_{40}}{{}_5L_{35}}$. This ability of the matrix S to "survive" the population follows from the convention of row by column multiplication. If we are interested in an operator which when applied to the original age distribution will carry it forward two periods, we multiply instead by a matrix with seven non-zero elements, located in the positions below those of the preceding matrix:

$$\mathbf{S}^2 = \begin{bmatrix} 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ .99463 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & .99618 & 0 & . & . & . & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & .98128 & 0 & 0 \end{bmatrix}$$

This can be verified by multiplying S by itself. The inner product of the third row of S by its first column gives the first element in the third row of S^2, i.e., $.99633 \times .99829 = .99463$. This meets the requirement that the premultiplication of a column vector consisting of the number of persons in successive five-year age groups by S^2 will carry the first age group into the third, the second into the fourth, etc., because the non-zero elements of S^2 are in the sub-subdiagonal; these elements are products of successive pairs of the non-zero elements of S. Each higher power has one less non-zero element; after seven multiplications we have the same 9×9 matrix, but with only one non-zero element remaining:

$$S^8 = \begin{bmatrix} 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ .96557 & 0 & 0 & . & . & . & 0 & 0 & 0 \end{bmatrix}$$

The effect of operator S^8 is to transfer the age group 0–4 in 1960 to 40–45 in the year 2000, multiplying it by .96204, the probability of surviving 40 years. Matrix multiplication here is arithmetically identical with aging a population on the usual life table method.

A MATRIX FOR FERTILITY

But the study of real populations must simultaneously take account of births as well as of deaths and aging. The ordinary method of population projection would start with the fact that the average number of children born to women of 15–19 years in 1960 was .0899, and it would project this age-specific birth rate by applying it to the average number of women 15–19 exposed over the five-year period. Specifically it would multiply .0899 by the average of the first two figures opposite 15–19 in Table 1, giving the product 666,000. So continuing we obtain column 6 of Table 1. All that remains to be done is to multiply 4,421,000, the total of this column, (1) by the factor .48807 which is the fraction which girl births constitute of total births in 1960, (2) by .97612 to allow for the deaths of children under five years of age between 1960 and 1965, (3) by 5 because we assume the annual rates of birth of 1960 will continue for the five years. The product of these three numbers is 2.38207 and applying this to 4,421,000 we obtain 10,531,000 as the projected female population 0–4 in 1965.

If we rearrange this calculation so as to be able to apply factors to the age distribution of 1960, we find that we must multiply the 8,249,000 girls 10–14 of that year by .1068; the 6,586,000 girls 15–19 by .4135, etc. This may be seen as a vector inner product obtainable by premultiplying the 1960 age distribution by B, a birth

$$B = \begin{bmatrix} 0 & 0 & .1068 & .4135 & . & . & . & . & .0195 \\ 0 & 0 & 0 & 0 & . & . & . & . & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

matrix which has only seven non-zero elements, all in the first row.

MORTALITY AND FERTILITY COMBINED

Since what we want is to find the effect of mortality and fertility simultaneously, it is convenient to add S and B to secure the single matrix M. We intend to premultiply the column vector K by M, i.e., $(S + B)K =$

$$MK = \begin{bmatrix} 0 & 0 & .1068 & .4135 & .5416 & .3686 & .2007 & .0862 & .0195 \\ .99633 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .99829 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .99789 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .99689 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .99606 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .99477 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .99253 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .98867 & 0 \end{bmatrix} \begin{bmatrix} 9991 \\ 9187 \\ 8249 \\ 6586 \\ 5528 \\ 5536 \\ 6103 \\ 6402 \\ 5924 \end{bmatrix}$$

and it may be verified arithmetically that this gives the same population for 1965 as we obtained in Table 1. What we have said of

S and B separately applies to M, i.e., it follows from the row-by-column convention of matrix multiplication that M corresponds to the projection operation because its first row generates the children in the first age group of the following time period, and its subdiagonal multiplies each age group by the proportion which survives and places the survivors in the succeeding age group. This also follows in a more general way from the distributive rule to which matrices are subject:

$$SK + BK = (S + B)K = (M)K$$

The only difference between this and ordinary algebra is that the commutative rule does not apply; in removing brackets or performing other manipulation, we must always keep the order of the factors unchanged.

POWERS OF THE MATRIX

By choosing as the top row not exactly the proportion having a child in the given age group but a kind of average of the proportions for two successive age groups, we reproduce the ordinary results of a population projection to any number of decimal places in which we may be interested. The utility of this way of expressing the projection depends on a computer being available. Our access to the University of Chicago's IBM 7094 has been through the symbolic language known as FORTRAN. The following program results in a printout of the first 21 powers of the matrix M and the product of each of these by the age distribution which represents the jumping-off point of the projection. The time required for doing this with six entirely distinct sets of data, making 126 multiplications of 9×9 matrices, as well as an equal number of matrix by vector multiplications, and including the time in which the machine compiled the FORTRAN program into machine language and wrote the results for 1,770 lines of printout, was 57 seconds. The brevity of the program is due to the fact that the computer has the capacity to repeat the same operation with different numbers, and FORTRAN has been designed to call on this capacity. Statement 105, for example, asks the machine to perform a multiplication of one row of the

TABLE 2. MATRIX TO POWERS UP TO 21 AND POPULATION PROJECTION.

```
      DIMENSION  AM(9,9),AM2(9,9),AM3(9,9),A(9),A2(9)
  1   READ INPUT TAPE 5,100,AM
100   FORMAT (9F9.8)
      WRITE OUTPUT TAPE 6,103,AM
      READ INPUT TAPE 5,20,A
 20   FORMAT (9F6.0)
      WRITE OUTPUT TAPE 6,22,A
      DO 3 I = 1,9
      DO 3 J = 1,9
  3   AM2(I,J)= AM(I,J)
      DO 200 N = 1,21
      DO 105 J = 1,9
      A2(I)= A2(I)+ AM2(J,I)* A(J)
      DO 105 J = 1,9
      DO 105 K = 1,9
105   AM3(I,J)= AM3(I,J)+ AM(I,K)*AM2(K,J)
      WRITE OUTPUT TAPE 6,22,A2
 22   FORMAT (9F13.0)
      WRITE OUTPUT TAPE 6, 104, N
104   FORMAT (X6HPOWER 12////)
      WRITE OUTPUT TAPE 6,103, AM3
103   FORMAT (X9F13.7)
      DO 200 I = 1,9
      AM2 (I,J)= AM3(I,J)
200   AM3(I,J)= 0
      GO TO 1
      END
```

matrix to the square power by one column of the original matrix; this line is the heart of the program. After performing for all the combinations of row and column of the original matrix, it goes on to repeat, this time putting the cube of the original matrix thus attained in place of the square which occupied the postmultiply position in the previous cycle. In this fashion the process continues to the power 21, which is to say—since each power represents a five-year evolution—for the century following the date to which the survivorship and fertility rates apply. In between raising the matrix to its successive powers the computer multiplies each power by the original (1960) age distribution of the population, and so tells us the age distribution that would appear at the end of successive five-year time intervals if the original fertility and survivorship rates are maintained.

In order to study the properties of the matrix as it approaches stability, it has been found convenient to square it repeatedly; this involves an even simpler FORTRAN program than we presented above. With the United States 1960 data we obtain for M^{64} the array

$$M^{64} = \begin{pmatrix} 153.67462 & 171.16562 & . & . & . & . & . & . & 2.70034 \\ 137.97100 & 153.67462 & . & . & . & . & . & . & 2.42440 \\ 124.11578 & 138.24242 & . & . & . & . & . & . & 2.18094 \\ 111.60715 & 124.31011 & . & . & . & . & . & . & 1.96114 \\ 100.25858 & 111.66989 & . & . & . & . & . & . & 1.76173 \\ 89.98900 & 100.23142 & . & . & . & . & . & . & 1.58127 \\ 80.66677 & 89.84810 & . & . & . & . & . & . & 1.41746 \\ 72.14744 & 80.35910 & . & . & . & . & . & . & 1.26776 \\ 64.27686 & 71.59276 & . & . & . & . & . & . & 1.12946 \end{pmatrix}$$

It is easy to satisfy ourselves of the stability of this operator which converts an arbitrary age distribution to the age distribution that would prevail in $5 \times 64 = 320$ years with the given rates of mortality and fertility. For in order to obtain any but the first row of cells of M^{65} we need only multiply a single cell of M^{64} by the appropriate survivorship factor. Thus the second element of the second row of M^{65} is given by the product of .99633 (taken from the matrix on p. 74 by $171.1656227 = 170.537445$. Dividing this by the second element of the second row of M^{64}, which is 153.6746197, we have 1.1097307. Using the ratios of the third element of the third row of M^{65} to M^{64} gives us again 1.1097307. It appears that the stability applies to at least seven significant digits.

APPLICATION TO A HUMP OF BIRTHS

Having calculated these powers of M we are now in a position to study the effects of a thousand additional girls under five in 1960. In fact, the computations are contained in the powers of M, and we can simply write down the results from the first columns of the several powers of the matrix as contained in the 7094 printout. These first columns would be the only non-zero components in the products M^tK where K consists in 1,000, with zeros below it. I present these figures at 25-year intervals for 1960 to 2060 (Table 3). It will be seen that in 1985 the original girls are 25–29 years of

TABLE 3. NUMBERS IN THE SEVERAL AGE GROUPS RESULTING FROM 1,000 GIRLS UNDER 5 IN 1960, AT FERILITY AND SURVIVORSHIP RATES OF THE UNITED STATES IN 1960.

Age	*1960*	*1985*	*2010*	*2035*	*2060*
0– 4	1,000	536	641	972	1,589
5– 9		409	535	846	1,412
10–14		106	364	710	1,249
15–19			281	633	1,127
20–24			371	634	1,044
25–29		986	528	632	958
30–34			402	526	833
35–39			103	356	695
40–44				273	613
Total	1,000	2,037	3,225	5,582	9,520

age, and the distribution of their descendants is non-overlapping with the distribution of their own ages—there would have been something wrong with the calculation if it had shown such overlapping, since mothers cannot be in the same five-year age group with their own daughters. Fifty years later, in the year 2010, the original babies have passed out of the reproductive ages, and the peak of their daughters is at 25–29. The distribution of the daughters, however, is not separate from that of their granddaughters; the two overlap in the trough which appears at 15–19. The granddaughters in the 0–4 group in 2010 are 641 in number. Examination of the distribution after 75 years shows the granddaughters having their (very slight) peak of numbers at age 20–24, which is 55 years younger than the original cohort of extra births, and a new peak (of great-granddaughters) coming up as indicated in the 972,000 children 0–4 in the year 2035. Since the length of the generation is about 25 years, each of the successive columns of the table represents a new generation if one follows along any horizontal line. By 2060 the overlapping process has gone far enough so that the several generations are represented by small inflections and no peaks at all.

By how much would this result have differed with the survivorship and fertility rates of 1940? For the United States of 1940,

$$M = \begin{pmatrix} 0 & 0 & .0633 & .2199 & .2987 & .2384 & .1498 & .0737 & .0203 \\ .99093 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .99589 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .99425 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .99093 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .98861 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .98640 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .98305 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .97799 & 0 \end{pmatrix};$$

again using the first columns of its powers multiplied by 1,000 we get Table 4 corresponding to the preceding Table 3. The differences in level between Tables 3 and 4 are spectacular. The reason for the differences, as may be seen by comparison of the top rows of the two matrices, is that at the principal ages of childbearing 1960 age-specific rates were nearly double those of 1940.

TABLE 4. NUMBERS IN THE SEVERAL AGE GROUPS RESULTING FROM 1,000 GIRLS UNDER 5 IN 1960, AT FERTILITY AND SURVIVORSHIP RATES OF THE UNITED STATES IN 1940.

Age	*1960*	*1985*	*2010*	*2035*	*2060*
0– 4	1,000	290	203	195	201
5– 9		214	171	186	198
10–14		62	150	187	199
15–19			166	199	201
20–24			227	209	197
25–29		961	279	196	188
30–34			205	164	178
35–39			58	141	177
40–44				154	185
Total	1,000	1,527	1,459	1,631	1,724

In both Tables 3 and 4 we are approaching the stable distribution by the end of the 100-year period. At the 1940 rates this stable distribution is nearly flat as one moves from age to age, but that on the 1960 rates shows less than half the number of women at age 40–44 as at 0–4. These results are an approximation to the ultimate or stable age distribution which arises from the given fertility and mortality rates. We can also make a comparison with a population which is increasing much more rapidly than that of the United States, and introduce the 1960 matrix for Mexico. We have been fortunate in securing the 1960 life tables prepared by Señora Zulma Recchini under the sponsorship of the Centro Latinoamericano de Demografía, and have combined these with the fertility and population figures given for the same year in the 1962 United Nations Demographic Yearbook. For the Mexican matrix and the 1,000

$$M = \begin{bmatrix} 0 & 0 & .1145 & .4413 & .6689 & .6384 & .5132 & .3197 & .1025 \\ .96495 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .99017 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .99234 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .98881 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .98106 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .98046 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .97649 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .97147 & 0 \end{bmatrix}$$

hypothetical extra population 0–4 in 1960 we have the calculations in Table 5. As before, each of the columns below is transformed into

TABLE 5. NUMBERS IN THE SEVERAL AGE GROUPS RESULTING FROM 1,000 GIRLS UNDER 5 IN 1960, AT FERTILITY AND SURVIVORSHIP RATES OF MEXICO IN 1960.

Age	*1960*	*1985*	*2010*	*2035*	*2060*
0– 4	1,000	627	1,001	2,235	5,245
5– 9		403	716	1,787	4,296
10–14		104	567	1,537	3,640
15–19			526	1,372	3,081
20–24			562	1,186	2,552
25–29			577	921	2,055
30–34		920	377	669	1,670
35–39			96	523	1,416
40–44				474	1,238
Total	1,000	2,054	4,422	10,704	25,193

the succeeding column by premultiplying it by M^5. Aside from the great difference in rates of increase, comparison of the United

States and Mexico shows a sharper separation between the generations in the former country. Between daughters and granddaughters of the original 1,000, for instance, as they appear in 2010 there is a trough at age 15–19 in the 1960 United States data, which is deeper both in relative and in absolute terms than that in the Mexican. This is related not so much to the level of fertility as to the shorter span of years over which women in the United States bear their children.

DESCENDANTS OF A GIVEN AGE GROUP

We draw attention to one other aspect of the survivorship-fertility matrices: the meaning of the individual cells of M^t. For concreteness in the exposition, consider the fourth cell in the eighth row of M^{15}, which is .590 with the United States 1960 data. This would be multiplied by the fourth age group of the original age distribution to secure the eighth age group 15 time periods (i.e., 75 years) later. In other words, 1,000 women 15–19 in 1960 would give rise to 590 women in the age group 35–39 in the year 2035 with the fertility-survivorship pattern of 1960. These 590 would not be the sole descendants of the original 1,000 of course, but only those who would be 35–39 in the given year; adding them to the other elements of the fourth column of the matrix $1{,}000M^{15}$ we secure the total of 7,271 for ages 0–45. It is not possible from this calculation to divide the 950 into grandchildren, great-grandchildren, etc., of the original 1,000; we could have kept the generations separate by placing successive sets of babies into ever-new matrices, but this would have greatly complicated the task. What the simple approach of this exposition provides is the statement that at the rates of survivorship and fertility of the United States in 1960, 1,000 girls aged 15–19 in 1950 would "yield" 590 women 35–39 through the several possible routes of descent, i.e., as granddaughters and great-granddaughters. At the United States 1940 rates the corresponding number is 195; at the Mexican 1960 rates it is 1,115.

To generalize this somewhat, we can find the descendants of the women (on a per woman basis) in the ith age group at the initial period who are in the jth group at the tth period by noting the ith element of the jth row of M^t. Similarly if we want to know the total (to age 45) of the descendants of the initial ith age group period we add the ith column of M^t.

CONCLUSION

Out of a short experience it is already apparent along what lines machine computation will alter demographic investigation. When the writer first set out to use the computer, he fed into it a set of birth rates along with the number-living column of a life table and read out the intrinsic rate of natural increase. Apart from this, he experimented on the construction of life tables, and also on the multiplying of matrices. He put in the elements of a matrix which had been calculated by hand and read out its latent roots. Each of these steps was preceded by long work at the desk calculator, working out the data for a short problem which would occupy the computer for a fraction of a second. In short, the computer was being used like a desk calculator, disregarding its ability to carry through a long sequence of calculations.[5]

The radical change was the discovery that it was in the long run easier to program the entire sequence and have the computer at each stage prepare the data for the next stage. In the present program of 800 FORTRAN statements, there are sub-programs for reading in and checking the raw data, which consist of a mere 90 or so figures copied out of primary sources or the United Nations Demographic Yearbook, without any processing whatever. The computer then calculates a life table (checking itself by the use of three different methods); works out the elements of the 9×9 matrix; takes the matrix to high powers; finds the intrinsic rate of increase by four methods whose agreement is the assurance that these many parts of the program are in agreement; uses this to work out the stable population; goes on to calculate moments and cumulants of the original population distribution by age, as well as of the life table functions and the maternity function. The sequence consists of a series of sub-routines which one by one fill the cells provided as receptacles for the several answers. It prints out some 800 lines of results in about 12 seconds. This sequence is a kind of skeleton on which other sub-routines can be hung, written by other workers in the field. Mr. E. M. Murphy has completed a sub-routine for finding the inverse of the fundamental matrix, and for taking it to high powers; he has also worked out a way of handling the 20×20 matrix that will embrace all ages instead of being confined to ages 0–44 as is the matrix of this article. Mr. J. Palmore has worked out a standardization routine for birth rates.

These will be tied into the existing sequence, starting from raw data, as was said. So far all this takes in only the female portion of the population, but plans are under way to do the same thing for the male side. Having within the same program the trajectories of the male and female populations, it ought to be possible to program marriage rates and the reciprocal changes by which the two sexes remain in balance. We believe that this and other kinds of simulation will be attainable in the near future.

REFERENCES

[1] Bogue, Donald J., THE POPULATION OF THE UNITED STATES, New York, The Free Press of Glencoe, 1959, Chapter 26, "Future Population."

[2] Bernardelli, H., Population Waves. *Journal of the Burma Research Society,* 31, 1–18, 1941.

[3] Leslie, P. H., On the Use of Matrices in Certain Population Mathematics, *Biometrika,* 33, 183–212, November, 1945.

[4] Lopez, Alvaro, PROBLEMS IN STABLE POPULATION THEORY, Princeton, N. J., Office of Population Research, 1961.

[5] The elements of the matrices S, B, and M cited in this paper were worked out on a desk calculator with existing life tables. It is hoped that similar results secured by a computer program, which works out the elements of the matrix in the first place from raw data giving births, deaths, and population by age, will be available soon. Because of rounding off, and for other reasons, it is not to be expected that the two sets will be identical.

ACKNOWLEDGMENT

The work on which this article is based was supported by a Ford Foundation grant to the Population Research and Training Center of the University of Chicago.