APPENDIX 2: MATRIX SOLUTION TO LINEAR EQUATIONS AND MARKOV CHAINS

Direct Solution and Convergence Method

Before computer programs offered ready solutions, problems such as Markov chains were solved in a direct manner, by algebraically manipulating the equations. This direct solution requires an understanding of simple matrix arithmetic, and very careful attention to calculating the numbers correctly. The convergence method is now easier, although it requires many more calculations. Without the computer, we would never even consider using this approach—with a computer, it is the best choice.

General Matrix Form

A matrix is a rectangular arrangement of elements into rows and columns. A matrix *A* is said to be *m x n* (pronounced “m by n”) if there are *m* rows and *n* columns in *A.*

$$a_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Certain properties of a matrix make it a valuable tool for solving simultaneous linear equations. These elementary matrix operations, called transformations, allow you to alter the rows (which will represent equations) without changing the solution. There are three basic row operations:

1. Multiplication or division of all elements of the row by any number.
2. Interchanging of any two rows (and consequently of all rows).
3. The addition or subtraction of the elements of one row with the corresponding elements of another.

To relate the matrix to simultaneous linear equations, consider a three-equation example, where only the coefficients, a, remain as unknowns:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = a_{14} \quad (1)$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = a_{24} \quad (2)$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = a_{34} \quad (3)$$

The three elementary row operations can now be interpreted in terms of simple operations on an equation:

1. When both sides of an equation are multiplied or divided by the same value, the results are equal.
2. Any two equations in a system of equations can be interchanged with no effect.
3. When two equals are added or subtracted, the results are equal.

Direct Solution

Putting these rules into use, write the coefficients of the three simultaneous linear equations as a 3 × 4 coefficient matrix,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$$

with the objective of reducing the matrix to the form

$$\begin{pmatrix} 1 & 0 & 0 & A_1 \\ 0 & 1 & 0 & A_2 \\ 0 & 0 & 1 & A_3 \end{pmatrix}$$

which would mean that $x_1 = A_1$, x_2, $= A_2$, and $x_3 = A_3$ since

$$1 \times x_1 + 0 \times x_2 + 0 \times x_3 = A_1$$

$$0 \times x_1 + 1 \times x_2 + 0 \times x_3 = A_2$$

$$0 \times x_1 + 0 \times x_2 + 1 \times x_3 = A_3$$

To achieve the results, divide the first equation by a_{1b} the first element, leaving

$$\begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & \frac{a_{13}}{a_{11}} & \frac{a_{14}}{a_{11}} \end{pmatrix} \qquad (1\text{-}1)$$

and then multiply by a_{21} to get the first elements in rows one and two the same:

$$\begin{pmatrix} a_{21} & \dfrac{a_{12} \times a_{21}}{a_{11}} & \dfrac{a_{13} \times a_{21}}{a_{11}} & \dfrac{a_{14} \times a_{21}}{a_{11}} \end{pmatrix} \qquad (1\text{-}2)$$

Now subtract (1-2) from (2) and get

$$\begin{pmatrix} 0 & a_{22} - \dfrac{a_{12} \times a_{21}}{a_{11}} & a_{23} - \dfrac{a_{13} \times a_{21}}{a_{11}} & a_{24} - \dfrac{a_{14} \times a_{21}}{a_{11}} \end{pmatrix} \qquad (1\text{-}3)$$

That successfully eliminates the first element a_{21} from the second equation. By going back and multiplying (1,1) by a_{31} and subtracting the resulting equation from (3), a_{31} can be eliminated from equation (3). Now column 1 looks like

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Column two or any column can be operated upon in the same manner as the first column:

1. Divide row n by the element in position n (for example, a_{nn}), thereby setting $a_{nn} = 1$.
2. Multiply row n by the corresponding element in row $i \neq n$, so that $a_{nn} \times a_{in} = a_{in}$.
3. Subtract row n from row i, resulting in $a_{in} = 0$, and all other elements reduced by the corresponding element in row n.

Continue this procedure for each row i until all elements

$$a_{1n}, a_{2n}, \ldots, a_{i\text{-}1n}, a_{1+1}n, \ldots, a_{mn} = 0 \quad \text{and} \quad a_{in} = 1$$

For example, beginning with

$$\begin{pmatrix} 2 & 4 & 8 & 34 \\ 6 & 5 & 2 & 22 \\ 3 & 6 & 5 & 30 \end{pmatrix}$$

Divide row 1 by the value 2 (position a_{11}):

$$\begin{pmatrix} 1 & 2 & 4 & 17 \\ 6 & 5 & 2 & 22 \\ 3 & 6 & 5 & 30 \end{pmatrix}$$

Multiply row 1 by the value 6 to get (6 12 24 102) and subtract the new calculated row from row 2:

$$\begin{pmatrix} 1 & 2 & 4 & 17 \\ 0 & -7 & -22 & -80 \\ 3 & 6 & 5 & 30 \end{pmatrix}$$

Multiply row 1 by the number 3 to get (3 6 12 51), then subtract the calculated row from row 3:

$$\begin{pmatrix} 1 & 2 & 4 & 17 \\ 0 & -7 & -22 & -80 \\ 0 & 0 & -7 & -21 \end{pmatrix}$$

Column 1 is now completed. Divide row 2 by –7 and get:

$$\begin{pmatrix} 1 & 2 & 4 & 17 \\ 0 & 1 & +22/7 & +80/7 \\ 0 & 0 & -7 & -21 \end{pmatrix}$$

Multiply row 2 by 2 and subtract the result (0 2 44/7 160/7) from row 1 to eliminate position 2 in the first row:

$$\begin{pmatrix} 1 & 0 & -16/7 & -41/7 \\ 0 & 1 & +22/7 & +80/7 \\ 0 & 0 & -7 & -21 \end{pmatrix}$$

Because the element in row 3 and column 2 is already 0, move to row 3. Divide row 3 by –7 and get

$$\begin{pmatrix} 1 & 0 & -16/7 & -41/7 \\ 0 & 1 & +22/7 & +80/7 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Multiply row 3 by –16/7 and subtract the result (0 0 –16/7 –48/7) from row 1:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & +22/7 & +80/7 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Multiply row 3 by +22/7 and subtract (0 0 +22/7 +66/7) from row 2:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

The results show that $x_1 = 1$, $x_2 = 2$, $x_3 = 3$.

There are other ways to reduce the matrix to the final representation, but this technique is well defined and lends itself to being programmed on a computer.

Solution to Weather Probabilities Expressed as a Markov Chain

Let $A = P(\text{clear})_{i+1} = P(\text{clear})\text{i}$

$B = P(\text{cloudy})_{i+1} = P(\text{cloudy})_i$

$C = P(\text{rainy})_{i+1} = P(\text{rainy})_i$

because when they converge all ith elements will equal $(i + 1)$th elements. The equations are

$$A = 0.7A + 0.2B + 0.2C \qquad (1)$$

$$B = 0.25A + 0.6B + 0.4C \qquad (2)$$

$$C = 0.05A + 0.2B + 0.4C \qquad (3)$$

In addition

$$A + B + C = 1(4)$$

To solve this system of equations using matrices, convert and add equation (3) to (4),

$$-0.3A + 0.2B + 0.2C = 0 \qquad (1')$$

$$0.25A - 0.4B + 0.4C = 0 \qquad (2')$$

$$1.05A + 1.2B + 0.4C = 1 \qquad (3')$$

which becomes the matrix

$$\begin{pmatrix} -.3 & .2 & .2 & 0 \\ .25 & -.4 & .4 & 0 \\ 1.05 & 1.2 & .4 & 1 \end{pmatrix} \begin{matrix} (1') \\ (2') \\ (3') \end{matrix}$$

The following are key steps in the solution:

1. Reduce the first row,

$$\begin{pmatrix} 1 & -.6667 & -.6667 & 0 \\ .25 & -.4 & .4 & 0 \\ 1.05 & 1.2 & .4 & 1 \end{pmatrix}$$

and make the leading entries of rows 2 and 3 zero:

$$\begin{pmatrix} 1 & -.6667 & -.6667 & 0 \\ 0 & -.2333 & .5667 & 0 \\ 0 & 1.9000 & 1.1000 & 1 \end{pmatrix}$$

2. Reduce the second row,

$$\begin{pmatrix} 1 & -.6667 & -.6667 & 0 \\ 0 & 1 & -2.4291 & 0 \\ 0 & 1.9000 & 1.1000 & 1 \end{pmatrix}$$

and make the second entries of rows 1 and 3 zero

$$\begin{pmatrix} 1 & 0 & 2.2857 & 0 \\ 0 & 1 & -2.4291 & 0 \\ 0 & 0 & 5.7143 & 1 \end{pmatrix}$$

3. Reduce the third row,

$$\begin{pmatrix} 1 & 0 & 2.2862 & 0 \\ 0 & 1 & -2.4291 & 0 \\ 0 & 0 & 1 & .1750 \end{pmatrix}$$

and make the third entry of rows 1 and 2 zero:

$$\begin{pmatrix} 1 & 0 & 0 & .4000 \\ 0 & 1 & 0 & .4250 \\ 0 & 0 & 1 & .1750 \end{pmatrix}$$

Then $A = .4000$, $B = .4250$, and $C = .1750$.

Computer Program Direct Solution

The following FORTRAN program accepts a system of 10 equations in 10 variables (unknowns) and applies the matrix method of solution. This is done using the method of Gaussian elimination, discussed in the previous section.

```
       PROGRAM MATRIX
C----  Matrix solution to simultaneous linear equations
C----  Copyright 1986 PJ Kaufman
       DIMENSION A(10,10), C(10)
       OPEN(6,FILE='PRN')
       WRITE(*,7000)
7000   FORMAT(' Enter matrix size (n)>'\)
       READ(*,5000)N
5000   FORMAT(BN,14)
       IF(N.GT.1O)STOP 'Matrix limited to 10 x 10¢
       WRITE(*,7001) (1,1=1,N)
7001   FORMAT(' Enter matrix elements row by row under headings
     +        ' do not include constant to right of
     +        7X,10(12,'-xxxxx'))
       DO 20 I = 1,N
```

```
          WRITE(*,7003)1
7003   FORMAT(' Row',12,':'\)
       READ(*,5003) (A(I,J), J=1,N)
5003   FORMAT(BN,10F8.0)
  20   CONTINUE
    WRITE(*,7004) (J,J=1,N)
    WRITE(6,7004) (J,J=1,N)
7004   F0RMAT(/' Input matrix is:'//9X,10(I2,'-col '))
       DO 30 I = 1,N
          WRITE(6,7005)I,(A(I,J),J=1,N)
  30      WRITE(*,7005)I,(A(I,J),J=1,N)
7005     FORMAT(/' Row',I2,':',10F8.3)
       WRITE(*,7008)(I,I=1,N)
7008   FORMAT(/' Enter constant vector under headings...'/
     +         1X,10(I1,'xxxxxx '))
        READ(*,5008)(C(I),I=1,N)
5008    FORMAT(BN,10F8.0)
        WRITE(*,7009)(C(I),I=1,N)
        WRITE(6,7009)(C(I),I=1,N)
7009    FORMAT(/' Constant vector is:'/10F8.3)
C-----  Process row by row (Gaussian Elimination)
        DO 100 I = 1,N
           DIV = A(I,I)
           DO 40 J = 1,N
  40          A(I,J) = A(I,J)/DIV
           C(I) = C(I)/DIV
C-----  Zero out column I for each row
           DO 60 J = 1,N
              IF(J.EQ.I)GOTO 60
              FACT0R = A(J,I)
              DO 50 K = 1,N
  50             A(J,K) = A(J,K) - A(I,K)*FACT0R
              C(J) = C(J) - C(I)*FACT0R
  60          CONTINUE
 100      CONTINUE
        WRITE(*,7007)(C(I),I = 1,N)
        WRITE(6,7007)(C(I),I = 1,N)
7007    FORMAT(/' Solution vector is: '/10F8.3)
        CALL EXIT
        END
```

Sample Computer Printout

```
MATRIX
Enter matrix size (N)>3
Enter matrix elements row by row under headings...
do not include constant to right of =
        1-xxxxx 2-xxxxx 3-xxxxx
Row 1:  2       4       8
Row 2:  6       5       2
Row 3:  3       6       5
Input matrix is:
        1-col       2-col       3-col
Row 1:  2.000       4.000       8.000
Row 2:  6.000       5.000       2.000
Row 3:  3.000       6.000       5.000
Enter constant vector under headings.,
lxxxxxx 2xxxxxx 3xxxxxx
34      22      30
Constant vector is:
  34.000  22.000  30.000
Solution vector is:
   1.000   2.000   3.000
```

Convergence Method

This method performs a series of matrix multiplications until the difference between the new and previous matrix is very small.[1] Following the procedure in Chapter 2, "Basic Concepts," we can create a 3 × 3 frequency matrix by counting the number of up, down, and neutral days that follow other up, down, and neutral days. We use the term *neutral* to allow very small price changes to be considered in this group, rather than limit it to only those days with zero changes. For this example we will look at the number of up, down, and neutral days that follow a 5-day trend that was considered up, down, or neutral on the previous day. Suppose the results were those shown in the *frequency matrix F*:

		Next Day Price Change			
		Up	**Neutral**	**Down**	**Total**
Previous day trend	Up	70	40	30	140
	Neutral	50	30	45	125
	Down	35	45	65	145

Divide each item in *F* by the total given at the end of that row, and get the probability of each occurrence in a *transition matrix T*.

		Next Day Price Change		
		Up	**Neutral**	**Down**
Previous day trend	Up	0.500	0.285	0.214
	Neutral	0.400	0.240	0.360
	Down	0.241	0.310	0.448

Now it is necessary to perform matrix multiplication. To multiply matrix *A* by matrix *B,* we multiply the corresponding items in row *i* of *A* by the corresponding item in column *j* of *B,* add those products together to get the item in row *i*, column *j* of the new matrix *C.* If we have two 3 × 3 matrices *A* and *B,* and we wanted to find the element in row 2, column 1 of the new product matrix *C*, we would multiply and add

$$c_{21} = a_{21} \times b_{11} + a_{22} \times b_{21} + a_{23} \times b_{31}$$

The general formula for this is

$$c_{ij} = \sum_{k=1}^{N} (a_{ik} \times b_{kj})$$

For spreadsheet users, there is no general formula to copy from one cell to another. For matrix *A* of 3 columns and 3 rows located in columns *A*, *B*, *C*, rows 1, 2, 3; matrix B in columns D, E, F, rows 1, 2, 3; and matrix C in rows G, H, I, columns 1, 2, 3, we enter the formula for c_{11} in cell $G1$ as

$$\text{cell } G1 = A1{*}D1 + B1{*}D2 + C1{*}D3$$

While this is clearly tedious, at least the arithmetic will be correct.

Iterative Matrix

We can now find the solution to the Markov chain, the long-term probabilities, by performing a series of matrix multiplications beginning with the transition matrix.

1. Multiply the transition matrix T by itself to get the first iterative matrix I_1,

$$I_1 = T \times T$$

2. Multiply the iterative matrix by the transition matrix to get the next iterative matrix,

$$I_2 = I_1 \times T$$

3. Continue to multiply the last iterative matrix by the transition matrix until the new iterative matrix is unchanged (or very close) to the previous iterative matrix,

$$I_n = I_{n-1} \times T$$

When each element satisfies the condition

$$\text{abs}(I_{ij}(n) - I_{ij}(n-1)) < 0.001$$

then the iterative matrix holds the final long-term probabilities of the Markov chain.

[1]George R. Arrington, "Markov Chains," *Technical Analysis of Stocks & Commodities* (December 1993).