Module 12: Matrices and Array Formulas – Applications in economics

---

Learning Goal for Module 12

Important models in economics feature matrices. Array formulas are a common method in Excel for calculating matrix operations, some of which are so common they are now functions – the **=SUMPRODUCT** being a common example. Microsoft 365 has added improved functionality in the form of dynamic arrays that reduce the work in setting up and using array formulas.

This Module examines models in economics that use matrices, such as the input-output model, which expresses linear regression in matrix format. In passing, this Module presents additional Excel procedures that will speed up your work.

By the end of this Module, you will be able to
- Perform basic matrix operations using Excel
- Name ranges for easier computation
- Express linear models such as input-output analysis and regression in matrix format
- Use common array functions
- Appreciate the power of dynamic arrays in simplifying Excel calculations

1. Introduction

A matrix is a rectangular array/grid/table/spreadsheet. In the broadest interpretation, a matrix can contain letters, words, symbols, etc., but for applications in economics, statistics, and science, matrices usually contain numbers. The dimension of a matrix, usually written r x c, describes its shape, which is important in basic algebraic operations. The corresponding designation of a matrix in Excel is an array. Therefore, array formulas are operations on matrices.

1.1. Basic operations using matrices

The simplest possible matrices are unidimensional. The basic rules of algebra also work for matrices.

$$1 + 3 = 4 \qquad\qquad\qquad … 1$$

$$5x = 10 \rightarrow x = (1/5) \bullet 10 = 2 \qquad\qquad\qquad … 2$$

In the second equation, we often just think of moving the 5 across the equation to divide into 10, but in fact the process involves creating an *inverse operator* (1/5) and multiplying it by 10. This

may seem trivial, but the reason for thinking about solving an equation this way will quickly become apparent.

Here are examples of higher dimensional matrices

$$\begin{pmatrix} 2 & -6 & 8 \end{pmatrix}, \begin{pmatrix} .9 \\ -3 \\ 45 \end{pmatrix}, \begin{pmatrix} .2 & 1.1 & -.7 \\ 2 & 9 & -13 \\ 5 & 0 & 1 \end{pmatrix}, \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

where the first and second are row and column matrices, respectively; the third is a rectangular (square) 3 x 3 matrix, and the last a generic m x n matrix where letters signify constants or unknowns as in any algebraic setting. The *order of a matrix* is its size, expressed as m x n. Row or column matrices often have the designation of *vectors*. Each number in a matrix is an *element.*

> **See Example 12A.xlsx**

1.2. Naming matrices

Matrix addition, subtraction, and scalar multiplication are straightforward, but before presenting details, here is how to name a range of cells.

1. Highlight the range.
2. Now on the Ribbon bar, click FORMULA → Define Name, and complete the process in the drop-down menu; notice how the range appears in the menu. You can choose whether the name applies to the Worksheet or the Workbook. (In general, choose the former, until you start working with complex models.)
3. You can name a vector in a row or column, or in an array.
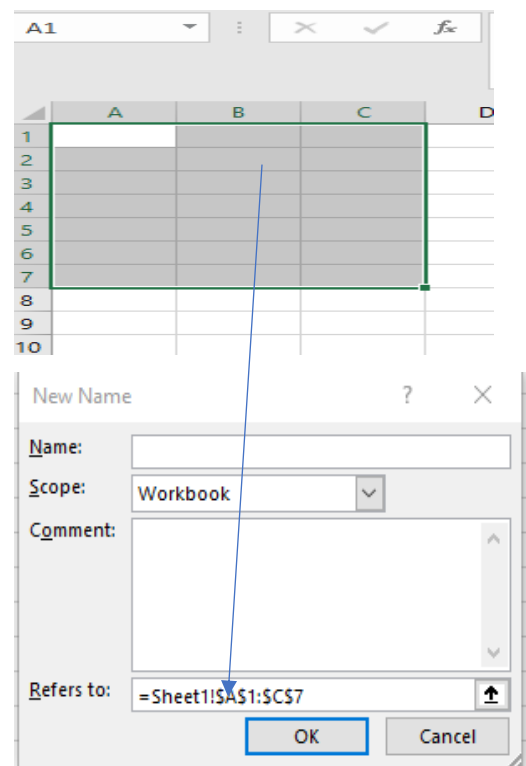
Module 12: Matrices and Array Formulas – Applications in economics

1.3. Matrix addition, subtraction, and multiplication - review

Matrices can be added, subtracted, multiplied, and divided, with the important condition that such operations apply only to *conformal matrices*. As shorthand, capital letters usually identify matrices.

$$Q = \begin{pmatrix} 200 & 150 & 50 \\ 300 & 250 & 300 \end{pmatrix} \quad P = \begin{pmatrix} 2.00 \\ 3.00 \\ 4.00 \end{pmatrix} \quad C = \begin{pmatrix} 1.75 \\ 2.25 \\ 3.75 \end{pmatrix}$$

$$A = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}, B = \begin{pmatrix} -1 & 6 \\ 4 & 2 \end{pmatrix}, C = \begin{pmatrix} 3 & -1 & 3 \end{pmatrix}$$

It is possible to add and subtract A and B, but not A and C or B and C. One can also multiply A and B, but one cannot execute A*C nor B*C (where A*C refers to A multiplied by C).

$$A - B = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} - \begin{pmatrix} -1 & 6 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 2-(-1) & 4-6 \\ 6-4 & 8-2 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ 2 & 6 \end{pmatrix}$$

Matrix multiplication follows specific rules, requiring the matrices to have conformal dimensions. Here are three matrices and their dimensions:

$$A = \begin{pmatrix} 2 & 4 & 8 \\ 1 & 2 & 4 \\ 3 & 6 & 12 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, C = \begin{pmatrix} 3 & 2 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 7 & 14 \\ -1 & -2 & -3 \\ 3 & 2 & 1 \end{pmatrix}$$

$$\quad\quad 3 \times 3 \quad\quad\quad 3 \times 1 \quad\quad 1 \times 3 \quad\quad\quad (3 \times 3)$$

The multiplication rule is that the "inner" dimensions must be the same. This means that only certain combinations can be multiplied. So, the following operations are defined.

$$AD = \begin{pmatrix} 22 & 22 & 24 \\ 11 & 11 & 12 \\ 33 & 33 & 36 \end{pmatrix}, B = (10), AB = \begin{pmatrix} 32 \\ 16 \\ 51 \end{pmatrix}$$

The mental math approach visualizes the following approach with the two 2x2 matrices A and B, multiplied to produce C:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \quad \dots \mathbf{3}$$

$$
\begin{aligned}
c11 &= a11 * b11 + a12 * b21 \\
c12 &= a11 * b12 + a12 * b22 \\
c21 &= a21 * b11 + a22 * b21 \\
c22 &= a21 * b12 + a22 * b22
\end{aligned}
\quad \dots \mathbf{4a\text{-}4d}
$$

Note that multiplying a 3x3 matrix times a 3x3 matrix produces a 3x3 matrix and multiplying a 3x3 matrix by a 3x1 matrix produces a 3x1 matrix. The operations BA and BD are **not** defined, but the operations CA and CD are, because the inner dimensions are the same.

$$
\begin{aligned}
BA &= (3 \times 3)(1 \times 3), \quad BD = (3 \times 3)(1 \times 3) \\
CA &= (1 \times 3)(3 \times 3), \quad CD = (1 \times 3)(3 \times 3)
\end{aligned}
$$

To manage the calculations in equations 4a – 4d, matrix multiplication uses an array formulation, which can be a bit tricky. Consider the two matrices from above, A and D. Applying equations 4a – 4d, now there are nine calculations because of the extra dimension.

Video: **Matrix Naming and Matrix Multiplication**

$$
A = \begin{pmatrix} 2 & 4 & 8 \\ 1 & 2 & 4 \\ 3 & 6 & 12 \end{pmatrix}, \; D = \begin{pmatrix} 1 & 7 & 14 \\ -1 & -2 & -3 \\ 3 & 2 & 1 \end{pmatrix}, AD = \begin{pmatrix} 22 & 22 & 24 \\ 11 & 11 & 12 \\ 33 & 33 & 36 \end{pmatrix}
$$

A useful operation is scalar multiplication, where a scalar constant λ pre-multiplies the entire array.

$$
\lambda A = \begin{pmatrix} 2 & 4 & 8 \\ 1 & 2 & 4 \\ 3 & 6 & 12 \end{pmatrix} = \begin{pmatrix} 2\lambda & 4\lambda & 8\lambda \\ 1\lambda & 2\lambda & 4\lambda \\ 3\lambda & 6\lambda & 12\lambda \end{pmatrix}
$$

Example: A store with four branches sells three types of tires, with the following table (matrix) showing the value of the inventory.

Tire Type→	Off Road	Winter	Summer
Branch↓			
St. Vital	$250,000	$45,000	$12,000
St. James	$120,000	$34,000	$100,000
East Kildonan	$50,000	$103,000	$98,000
Transcona	$102,000	$96,000	$32,000

Module 12: Matrices and Array Formulas – Applications in economics

---

If the store introduces a 15% price reduction to tire type 1, a 7% reduction for tire type 2, and a 2% reduction for tire type 3, by how much does the total inventory lose value? (Answer: $102,600. *HINT – use the =SUMPRODUCT, which is an array formula embedded in a regular formula*.)

Example: Create a dashboard so that the discounts can vary to show the implications of policy.].

> See Example 12B.xlsx

1.4. Matrix multiplication

Matrix multiplication follows the Distributive Law. Imagine a store sells three commodities in two branches, with the selling prices and costs as noted. The goal is to calculate profits and then change prices to recalculate profits.

Quantities Sold (Q)			
	G1	G2	G3
Branch 1	200	150	50
Branch 2	300	250	300

Selling prices ($/unit) (P)	
G1	2.00
G2	3.00
G3	4.00

Costs ($/unit) (C)	
G1	1.75
G2	2.25
G3	3.75

Develop this in matrix format:

$$Q=\begin{pmatrix} 200 & 150 & 50 \\ 300 & 250 & 300 \end{pmatrix} \quad P=\begin{pmatrix} 2.00 \\ 3.00 \\ 4.00 \end{pmatrix} \quad C=\begin{pmatrix} 1.75 \\ 2.25 \\ 3.75 \end{pmatrix}$$

From microeconomic theory, Total Revenue (TR) equals price x quantity, and Total Costs (TC) equals cost x quantity, or

$$TR=PxQ \text{ and } TC=CxQ$$

However, according to the matrix multiplication rules, we need to write *TR* as *QP* and *TC* as *QC.*

Example: Using the above information, calculate the profit.

$$TR=QP=\begin{pmatrix} 200 & 150 & 50 \\ 300 & 250 & 300 \end{pmatrix}\begin{pmatrix} 2.00 \\ 3.00 \\ 4.00 \end{pmatrix} \quad ;TC=QC=\begin{pmatrix} 200 & 150 & 50 \\ 300 & 250 & 300 \end{pmatrix}\begin{pmatrix} 1.75 \\ 2.25 \\ 2.75 \end{pmatrix} \text{ ...5}$$

Profit = TR - TC
Profit = $512.50

> See Example 12C.xlsx

This is not interesting but consider how the problem could become more complex.

> **Example 12D:** Assume the price elasticity of Good 1 is -.7, the price elasticity of Good 2 is -1.5, and the price elasticity of Good 3 is -.9. Recall that price elasticity predicts the change in quantity demanded in response to a change in price. We start from the existing data, assuming the prices and quantities above reflect an equilibrium.
>
> ε = Percent Change in Quantity/Percent Change in Price
>
> In each case, we use the given elasticity to calculate the quantity response to a change in price. This creates a change in total revenue that changes profit. Create a dashboard to model the effects of price changes on profit.

> **See Example 12D.xlsx**

2. Solving equations – The identity matrix

Imagine the average salaries in Canada for 2000, 2005, 2010, and 2015 are $20,000, $20,300, $20,600, and $20,900. Based on this, what is the expected salary in 2020? This is a simple l exercise, but it supports the demonstration of how to solve a system of linear equations. This leads to more substantive applications of matrix multiplication, such as linear regression and input-output analysis.

Here the dependent variable is salary, and the independent variable is time. The solution requires solving a set of linear equations, expressed as the general equation, $Y = a_0 + a_1 X$, where Y is salary and X is time. We could measure time directly as 2000, 2005, …, but we also let X be 0, 1, 2…, where every increment represents five years.

The equation system appears as follows.

$$a_0 + a_1 \cdot 0 = 20000$$
$$a_0 + a_1 \cdot 1 = 20300$$
$$a_0 + a_1 \cdot 2 = 20600$$
$$a_0 + a_1 \cdot 3 = 20900$$

Intuitively, the solution is a straight line (every five years is associated with a $3,000 increment in income). We use any two of these equations to find $a_0 = 3$ and $a_1 = 5$.

The identity matrix becomes the key to solving the general system of linear equations. Think of the simple problem at the start of this chapter, 5x=10. To solve this, we commonly just divide both sides by 5, but as noted, this is the same as multiplying both sides by 1/5 or the inverse of 5. Write this as

$$1/5 \cdot X = 1/5 \cdot 10$$
$$X = 2$$

Module 12: Matrices and Array Formulas – Applications in economics

The identity matrix for square matrices of different sizes appears as

$$\left(a_1\right) = \left(a_1\right)\left(a_1\right)^{-1} = (1), \quad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} a11 & a12 \\ a21 & a22 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The inverse of a square matrix is **not** $\begin{pmatrix} 1/a_{11} & 1/a_{21} \\ 1/a_{12} & 1/a_{22} \end{pmatrix}$, but each element a_{ij} divided by a special value, known as the determinant of A. Fortunately, Excel supplies a special array formula for calculating inverses.

The general expression of a system of linear equations appears as

$$AX = B,$$

where A is a matrix that can have any dimension. In the simple example, A is a 1x1 matrix, where A = 5 and A inverse (1/5) appears as A^{-1}, since $5^{-1} = 1/5$. What we need is a way to create A^{-1} for higher dimensional matrices for a system of equations.

$$\begin{matrix} 3x - 2y + z = 7 \\ 2x + y + 3z = 5 \\ x - y + 2z = 8 \end{matrix} = \begin{pmatrix} 3 & -2 & 1 \\ 2 & 1 & 3 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \\ 8 \end{pmatrix} = AX = B; \quad A = \begin{pmatrix} 3 & -2 & 1 \\ 2 & 1 & 3 \\ 1 & -1 & 2 \end{pmatrix} \text{ ... } 6$$

The inverse matrix A^{-1} is that matrix, when multiplied by A produces the identity matrix. Back to the simple example, the inverse matrix to A=5 (a 1x1 matrix) is A^{-1} = 1/5, since AA^{-1}=1. So, for the above system, we need to find A^{-1} for the matrix.

Module 12: Matrices and Array Formulas – Applications in economics

$$A = \begin{pmatrix} 3 & -2 & 1 \\ 2 & 1 & 3 \\ 1 & -1 & 2 \end{pmatrix}$$

The Excel array formula MINVERSE does that. It has a similar set-up as the MMULT formula, and by the { } you see this is an array formula (use CTRL-SHIFT-Enter and not the checkmark).

Let us make sure this is the inverse by multiplying A times A^{-1}.

A5		:	✕	✓	f_x	{=MINVERSE(A1:C3)}

	A	B	C	D	E	F
1	3	-2	1			
2	2	1	3		A	
3	1	-1	2			
4						
5	0.357143	0.214286	-0.5		A^{-1}	
6	-0.07143	0.357143	-0.5			
7	-0.21429	0.071429	0.5			
8						

In general AA^{-1} = I will be the identity matrix for square matrices. For the 3x3 matrix $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

See Example 12E.xlsx

See Video: Matrix Inverse

A9		:	✕	✓	f_x	{=MMULT(A1:C3,A5:C7)}

	A	B	C	D	E	F
1	3	-2	1			
2	2	1	3			
3	1	-1	2			
4						
5	0.357143	0.214285714	-0.5			
6	-0.07143	0.357142857	-0.5			
7	-0.21429	0.071428571	0.5			
8						
9	1	-5.55112E-17	-2.22045E-16			
10	1.11E-16	1	-4.44089E-16			
11	0	2.77556E-17	1			
12						
13						

What????

This is an example of round-off error. Use Format Cells to format these cells as numbers with no decimal and the 3x3 identity matrix will emerge. Practice with any size matrix.

©Gregory Mason, July 2022

If you see cells filled with #NUM!, that usually means an inverse does not exist. This occurs when the determinant of A does not exist, which usually reflects a situation when rows or columns are multiples of each other. If you can multiply a row (or column) by a number and add or subtract it from another row (or column) to get 0, an inverse to that matrix will not exist. This is the rule that says, in a linear system of k variables, you need k equations to find a unique solution (see equation 6 above).

The solution to 6 requires multiplying both sides of the equation by A^{-1}.

$$A^{-1}AX = A^{-1}B = \begin{pmatrix} .36 & .21 & -.5 \\ -.07 & .36 & -.5 \\ -.21 & .07 & .5 \end{pmatrix} \begin{pmatrix} 3 & -2 & 1 \\ 2 & 1 & 3 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ x \end{pmatrix} = \begin{pmatrix} .36 & .21 & -.5 \\ -.07 & .36 & -.5 \\ -.21 & .07 & .5 \end{pmatrix} \begin{pmatrix} 7 \\ 5 \\ 8 \end{pmatrix}$$

$$\qquad\qquad (3\times3) \qquad\quad (3\times3) \quad (3\times1) \qquad\quad (3\times3) \qquad\quad (3\times1)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -.43 \\ -2.69 \\ 2.88 \end{pmatrix} \text{ ... } \mathbf{7}$$

3. The regression models

A linear regression is a set of n (the number of observations) linear equations, each with k variables. The matrix format for linear regression imposes a standard "flat file" data structure, expressed as a rectangular matrix, with n rows and k columns. The data matrix X appears in Excel, with the first column enumerating the unit of analysis – dates for time series data and observation units (households, firms, countries, survey respondent numbers…) for cross-sectional data. Pooled time series/cross-section data follow the same rules with a slightly more complex file naming strategy.

Module 12: Matrices and Array Formulas – Applications in economics

Cross--sectional file of families, where first column, FAMNUM enumerates the record for each family in the data. The other columns show the variables defined in the data.

$$X = \begin{pmatrix} 1 & 0 & 143 & \dots \\ 1 & 0 & 143 & \dots \\ 1 & 0 & 155 & \dots \\ 1 & 62 & 277 & \dots \\ .. & .. & .. & \dots \\ 0 & 0 & 238 & \dots \end{pmatrix}$$

$$X = \begin{pmatrix} 657758 & 11360 & 27159 & \dots \\ 648066 & 11409 & 27559 & \dots \\ 661525 & 11722 & 27810 & \dots \\ 683842 & 12192 & 29130 & \dots \\ & \dots & & \\ & \dots & & \end{pmatrix}$$

|	A	B	C	D	E	F	G	H	I
1	Date	GDP_CA	GDP_NF	GDP_MB	GDP_BC	CAPI_CA	CAPI_NF	CAPI_MB	CAPI_BC
2	1981	657758	11360	27159	82793	174762	2450	4778	22357
3	1982	648066	11409	27559	79884	155802	2835	3880	18035
4	1983	661525	11722	27810	80054	154819	3153	4429	17698
5	1984	683842	12192	29130					
6	1985	715521	12605	30474					
7	1986	736324	12922	31095					
8	1987	759716	13365	31619					
9	1988	788475	14026	32013					
10	1989	813286	14376	32529					
11	1990	828839	14540	33011					
12	1991	827908	14454	32663					
13	1992	838885	14605	32676	107469	200381	3050	5502	28244
14	1993	849342	14544	32918	110549	198054	3545	5479	28462

Time--series file of four jurisdictions - CA, NF, MB, and BC families, where first column is Date (year), and the other columns show the variables defined in the data.

Note that the variable labels are not part of the data matrix, although most statistical programs, including Excel, will accommodate the inclusion of labels in the input. In time series data, time may be an independent variable but never as a raw date; the researcher usually creates a new time variable, where the first value becomes 1 and subsequent values are numbered consecutively. In the example here, 1981 = 1, 1982 = 2, and so on. The important idea is that these are n x k (n by k) matrices, designated by X. Recall the regression figure from Module 7 (Section 2.2). Each observation has a unique equation; for example:

$$y_i = a_o + a_1 X_{1i} + e_i$$

Module 12: Matrices and Array Formulas – Applications in economics
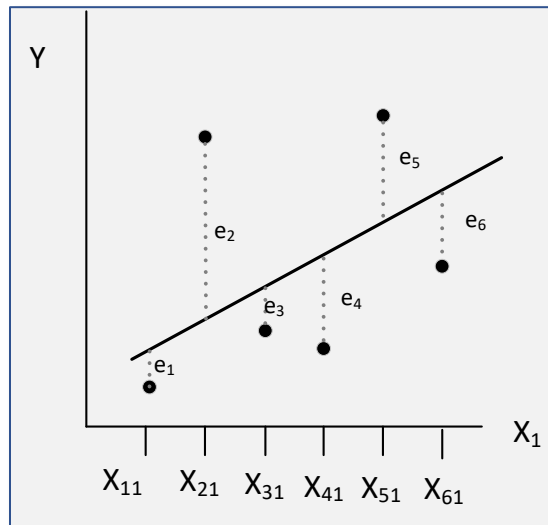
The entire model has six equations

$$y_1 = a_o + a_1 X_{11} + e_1$$
$$y_2 = a_0 + a_1 X_{21} + e_2$$
$$y_3 = a_0 + a_1 X_{31} + e_3$$
$$y_4 = a_0 + a_1 X_{41} + e_4 \qquad \text{… 8}$$
$$y_5 = a_0 + a_1 X_{51} + e_5$$
$$y_6 = a_0 + a_1 X_{61} + e_6$$

The independent variable X_{ij} has two subscripts to accommodate the multiple regression model.

Therefore, X_{31} is the third observation of variable X_1. In matrix form, the equations in 8 become:

$$
\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y6 \end{pmatrix}
=
\begin{pmatrix} 1 & x_{11} \\ 1 & x_{21} \\ 1 & x_{31} \\ 1 & x_{41} \\ 1 & x_{51} \\ 1 & x_{61} \end{pmatrix}
\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}
+
\begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{pmatrix}
\qquad \text{… 9}
$$

$$\begin{array}{cccc} (6x1) & (6x2) & (2x1) & (6x1) \end{array}$$

It is useful to check conformal rules (the brackets below each matrix). On the left-hand side, a 6x1 matrix is equal to a 6x1 plus a 6x1 matrix. Equation 9 becomes compact when written as:

$$Y = XA + e ,$$

where Y is an n x 1 vector (matrix), X is an n x 2 matrix, and e is an n x 1 column vector. (Note: Any matrix that has one of its dimensions =1 is often referred to as a "vector.") A multiple regression with these variables appears as

$$
\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix}
=
\begin{pmatrix} 1 & x_{11} & x_{12} & x_{31} \\ 1 & x_{21} & x_{22} & x_{32} \\ 1 & x_{31} & x_{31} & x_{33} \\ 1 & x_{41} & x_{42} & x_{43} \\ 1 & x_{51} & x_{52} & x_{53} \\ 1 & x_{61} & x_{62} & x_{63} \end{pmatrix}
\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}
+
\begin{pmatrix} e1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{pmatrix}
\qquad \text{…10}
$$

$$\begin{array}{cccc} (6x1) & (6x4) & (4x1) & (6x1) \end{array}$$

Module 12: Matrices and Array Formulas – Applications in economics

Mathematical manipulation (available in any econometrics or statistical text on regression) provides the estimate of a_i.

$$A = \begin{pmatrix} a_1 \\ a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_k \end{pmatrix} = (X'X)^{-1}X'Y \quad , \qquad \qquad \textbf{... 11}$$

where X' is the transpose of X, and $X'X^{-1}$ is the inverse of the $X'X$ matrix, also known as the cross-product matrix.

$$\text{If the matrix } C = \begin{pmatrix} 1 & 7 \\ 3 & 1 \\ -1 & -2 \\ 8 & 4 \end{pmatrix}, \text{ then } C' = \begin{pmatrix} 1 & 3 & -1 & 9 \\ 7 & 1 & -2 & 4 \end{pmatrix}, \text{ then}$$

$$CC' = \begin{pmatrix} 50 & 10 & 3 & 35 \\ 10 & 10 & -5 & 31 \\ -15 & -5 & 5 & -18 \\ 36 & 28 & -16 & 88 \end{pmatrix} \text{ and } C'C = \begin{pmatrix} 83 & 48 \\ 44 & 68 \end{pmatrix}$$

CC' is a 4x2 matrix times a 2x4 matrix creating a 4x4 matrix, and C'C is a 2x4 matrix times a 4x2 matrix creating a 2x2 matrix. In the single variable regression, a cross-product matrix appears as

$$X'X = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ x_{11} & x_{21} & x_{31} & x_{41} & x_{51} & x_{61} \end{pmatrix} \begin{pmatrix} 1 & x_{11} \\ 1 & x_{21} \\ 1 & x_{31} \\ 1 & x_{41} \\ 1 & x_{51} \\ 1 & x_{61} \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^{n} x_{i1} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 \end{pmatrix}$$

The next step is to calculate X'Y

$$X'Y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ x_{11} & x_{21} & x_{31} & x_{41} & x_{51} & x_{61} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i 1 y_i \end{pmatrix}$$

Module 12: Matrices and Array Formulas – Applications in economics

Example 12F: Given the following data, compute the regression coefficients for the linear model:

y	x
23	6
12	3
15	4
5	1
24	6
45	9
17	4

See Example 12F.xlsx

4. Input-output analysis[1]

Economists have always sought to create abstract and simplified models as a way of capturing the essence of economic processes. Viewing an economy as a set of inputs transformed into a set of outputs captures a basic reality. Quesnay's *Tableau Economique* (Economic Picture) represents an early attempt to see the economy as a continuous and circular flow of inputs transforming into outputs. The modern input-output table stems from the work of Wassily Leontief.

The concept starts with observations and assumptions, the most obvious being that any economy comprises multiple sectors, each supplying some sectors and dependent on others for inputs. This interrelationship and mutual dependency form the "structure" of the economy, expressed as a matrix showing the flows among the sectors. This matrix expression supports the analysis of exogenous shocks to the economy. Input-output (I-O) analysis often appears in economic impact analysis for large national investment projects involving government support when planners wish to calculate an overall return to public investment. Statements such that the economy will expand by $10 billion with a $500 million investment often have an I-O analysis supporting part of the argument.

Consider the following "economy" with three inputs Y, Z, and R that support three outputs Y, Z, and D, respectively. This unique economy has important assumptions:

I-O Flows		Outputs			$m
		Y	Z	D	Total
Inputs (Industry)	Y	420	150	270	840
	Z	84	200	216	500
	R	336	150		486
	Total	840	500	486	

1. The final demand for each community (row and column totals) are exogenous (determined by actors outside the economy).
2. In this simple model, each industry has one product.
3. The technology features constant returns to scale; doubling the input doubles the output, etc.

---

[1] This section draws from Rosser, M., and Lis, Piotr orter (2016) *Basic Mathematics for Economists*, London, Routledge.

4. This is a simple closed economy, and an industry may consume its own output (such as Z supplying itself 200 units in the example).
5. No surplus or profit exist – the sum of each row equals each column.

This is a simple general equilibrium model. Of course, we can relax some assumptions – not the one about constant returns to scale – and matrix algebra can support the use of this model to support important policy analysis.

Imagine an economy with two industries, one resource and the other as inputs. Official statistics agencies such as Statistics Canada typically manage the creation of input-output tables because the data requires detailed survey information on what firms purchase from whom and sell to whom. Typically, only central statistical agencies can gain the cooperation of enough firms to collect the needed data, comprising sales or purchases per period in currency. The entries in the table are financial flows per period, typically a year for national tables.

Referring to the example above, the total output of Industry Y is $840, $420 of which it consumes, while Industry Z consumes $150 and $270 is purchased in final demand – D (general consumers). The final demand category is important because changes in demand are what propels the model. The first column shows Sector Y using $420 of its own output, $84 from Sector Z, and $336 units of resource R for a total $840 in inputs.

This table of sales and consumption flow supports an input/output coefficient table that forms the core of I-O analysis. Each element of this matrix reflects the ratio of input to output flows. The numbers in brackets only show the calculation not included in the matrix.

I-O Coefficient Matrix	Y	Z
Y	.5	.3
Z	.1	.4
R	.4	.3

In the first upper left cell, .5 (420/840) is the proportion of total output from Industry Y that it uses in satisfying its own production needs. To illustrate, the production of electricity requires electricity. One unit of production of Y requires .5 units of its own product, .1 units of Industry Z's output, and .4 units of resource R. Note that the ratios of each column sum to one. This "technology" matrix (another name for the I-O matrix) shows how the input sector's outputs plus resources flow among the input sector.

The last column D (final demand) determines the input flows. To illustrate, with the I-O matrix above and assuming 600 units of R is available, is it feasible for final demand to be 324 of A and 261 of B? To answer this, use the I-O matrix to write two demand equations for industries A and B:

$$Y = .5Y + .30Z + 324 \qquad \ldots 12$$
$$Z = .1Y + .4Z + 261 \qquad \ldots 13$$

So,

$$.5Y = .3Z + 324 \quad \text{(multiply by 2)} \qquad \ldots 14$$
$$Y = .6Z + 648 \qquad \ldots 15$$

Substitute 16 into 14 to obtain $ZB=603.3$. Now substituting this into 14 gives Y = 1,010. Now, using the last row of the I-O matrix:

$$R=.4Y+.3Z=.4(1010)+.3(603.3)=404+181=585 \quad \text{... 16}$$

This is below the 600 units of R hypothesized, so this production is feasible. Let us set this up in matrices and show how to explore variations in the technology matrix to solve production problems.

Translating the current problem into a matrix formulation is straightforward.

$$Y=.5Y+.3Z+324$$
$$Z=.1Y+.4Z+261$$

$$(1-.5)Y-.3Z=324$$
$$-.1Y+(1-.4)Z=261$$

Let $A=\begin{pmatrix} 0.5 & 0.3 \\ .1 & .4 \end{pmatrix}$, and the final demand vector is $d=\begin{pmatrix} 324 \\ 261 \end{pmatrix}$, which creates the simultaneous system

> **See: Example 12G.xlsx**

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}-\begin{pmatrix} 0.5 & 0.3 \\ 0.1 & 0.4 \end{pmatrix}\right]\begin{bmatrix} Y \\ Z \end{bmatrix}=\begin{bmatrix} 324 \\ 261 \end{bmatrix}$$
$$=(I-A)x=d$$
$$x=(I-A)^{-1}d \quad\quad\quad \text{... 17}$$

Now our problem requires a solution using matrices. The first step is to compute $(I-A)^{-1}$ multiplied by d to produce the required values of Y and Z.

Example 12H: A central statistical agency collects the following information on sales and purchases in a small economy with two sectors X and Y, plus resources R. Remember an I-O model represents a closed economy, so everything produced is consumed. In this example, total demand = total output = 1152.

	X	Y	D	Total Output
X	1095	495	600	2190
Y	438	660	552	1650
R	657	495	0	1152
Total Demand	2190	1650	1152	

In this example, total final demand D = total resource consumption R. Sector X sends some of its output to itself (1095), some to Sector Y (495), and some to final demand (consumers) at 600. Sector Y sends some of its output to Sector X (438), some to itself (660), and some to final demand (552). Resources used to meet the demand of Sector X and Sector Y are 657 and 495 respectively. You can see that adding up (the more heavily shaded cells) shows that total output = total input = 1152.

Module 12: Matrices and Array Formulas – Applications in economics

The next step is to calculate the I-O coefficient table.

Calculate the I-O coefficient table (the A matrix) and estimate what happens if the demand for X increases to 690. Constructing the I-O matrix creates the relevant equations:

0.5	0.3
0.2	0.4
0.3	0.3

1095/2190 = .5
and so on.

$$X = .5X + .3Y + 690$$
$$Y = .2X + .4Y + 552$$
$$=$$
$$(1-.5)X - .3Y = 690$$
$$.2x + (1-.4)Y = 552$$
$$=$$

It is a coincidence that the A matrix is the same as the first two rows of the I-O matrix.

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] - \left[\begin{pmatrix} .5 & .3 \\ .2 & .4 \end{pmatrix} \right] \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 690 \\ 552 \end{bmatrix}$$

Example 12I. I-O problems are scalable, which is fortunate since actual I-O tables may have a hundred sectors and many types of resources.

	Agricultural	Manufacturing	Services	Demand	Total Inputs
Agricultural	996	495	325	700	2516
Manufacturing	438	490	300	585	1813
Services	657	495	570	305	2027
Resources	425	333	832		
Total Demand	2516	1813	2027		6356

The I-O matrix appears as

996/2516 = .4 and
so on.

0.40	0.27	0.16
0.17	0.27	0.15
0.26	0.27	0.28
0.17	0.18	0.41

$$A = .40A + .27M + .16S + 700$$

$$M = .17A + .27M + .15S + 585$$

$$S = .26A + .27M + .28S + 305$$

$$=$$

$$(1 - .40)A - .27M - .16 = 700$$

$$.17A + (1 - .27)M - .15S = 585$$

$$.26A - .27M + (1 - .28)S = 305$$

$$=$$

$$\left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} .40 & .27 & .16 \\ .17 & .27 & .15 \\ .26 & .27 & .28 \end{pmatrix} \right] \begin{pmatrix} A \\ M \\ S \end{pmatrix} = \begin{pmatrix} 700 \\ 585 \\ 305 \end{pmatrix}$$

See: Example 12H.xlsx

An I-O table represents a closed economy (no trade) with constant returns to scale. A common use for an I-O model is to simulate the impact of an endogenous expansion of final demand in one sector on the final demands for other sectors. Often proponents of large capital projects in one area of the country will argue that, because the project will use inputs from all parts of the country, the benefits are not local and will permeate the entire economy. Example 12I shows how to simulate the impact a change in final demand for one sector has on all sectors.

It is important to accept the limitations of I-O models. The assumption of constant returns to scale implies that, as industries expand, unit costs remain the same. Technical change does not enter in an I-O model without completely redoing the data collection and creating an entirely new model. The inputs for a project, such as a wind farm, must link to industries in the I-O table. If the table is older, it is possible it will not reflect current technology. No inflationary pressures exist, and analysis occurs at a point in time. The costs of financing new projects do not appear in the model.

Yet this model offers a way of understanding the impact of increases in demand in one sector on the entire economy. It allows one to assess the claim that a wind farm in BC will benefit industries in other

provinces. Of course, if the inputs for constructing a wind farm come from offshore, the I-O analysis will show only benefits for the rest of the industry.

On a final note, I-O will technically fail if it is not possible to create an inverse to the matrix (I-A). If a row (or column) is a linear function of other rows (columns) an inverse will not exist. If you develop an I-O model and the inverse function fails, fixes exist, but these are beyond the scope of this text.

5. Tables and dynamic arrays

These two topics do not deal with economic computations, but economists will need to access and manipulate information to create reports. Also, dynamic arrays update reports whenever source data change.

In 2018, Microsoft introduced the idea of a dynamic array. This, along with a new tables format, reduces the work needed to update reports as new data arrive and reduces the trickiness and tedium of developing array formulas. Dynamic arrays are only available in MS 365 and Excel 2021, but it is still wise to know the "old fashioned" way of doing things since many workplaces have legacy versions.

5.1. Tables

One of the quickest ways to increase processing speed in summarizing data is to place it in a table. The data used for regression in Module 6 were in rectangular arrays, with variable names as headers. Using the Sort functions in Excel under the Data tab will accomplish basic reorganizations of the information (arranging in ascending or descending order). Defining an array as a Table in Excel opens possibilities for quick reorganization.
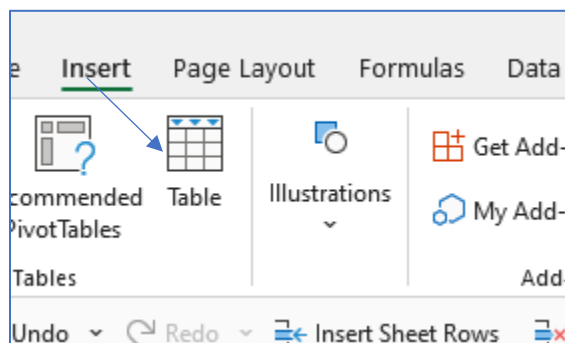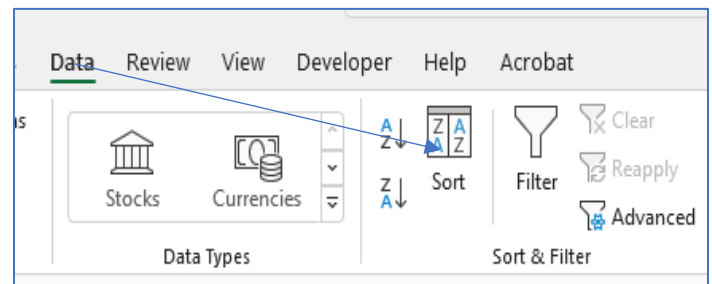
Selecting Insert → Table will produce the following menu, but it is much easier to view the video than read about the actions.
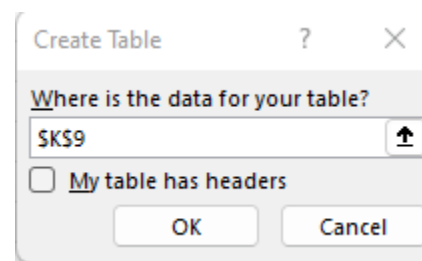
Video: Inserting and Naming Tables

See: Example 12II.xlsx

Module 12: Matrices and Array Formulas – Applications in economics

On the left is the data array with headers, while on the right is the same information inserted as an Excel Table. Again, the video illustrates the potential of this operation. Note the addition of drop-down menus on each header in an Excel Table.

	A	B	C	D	E
1	Name	Country	Sales	Age	
2	Zidong	USA	$99,771	44	
3	Zhipeng	Iceland	$60,823	34	
4	YuZhe	Cnina	$94,603	34	
5	Yubo	Ireland	$28,057	26	
6	Yi Fan	Cnina	$20,626	39	
7	Xinyue	Canada	$15,078	41	
8	Xiaowei	England	$4,509	28	
9	Xiao	Canada	$93,291	31	
10	Xiao	USA	$80,914	36	
11	Wenyue	Canada	$67,755	21	
12	Tatsuya	England	$51,324	25	
13	Syeed	USA	$74,144	31	
14	Peikun	Canada	$41,322	34	
15	Michaila	Canada	$36,409	29	
16	Larissa	Ireland	$89,600	36	
17	Kevin	USA	$61,515	23	
18	Junhyeok	Cnina	$65,624	32	
19	Joel	Canada	$65,257	41	
20	Jielin	England	$82,360	34	

	A	B	C	D
	Name	Country	Sales	Age
	Zidong	USA	$99,771	36
	Zhipeng	Iceland	$60,823	21
	YuZhe	Cnina	$94,603	25
	Yubo	Ireland	$28,057	27
	Yi Fan	Cnina	$20,626	36
	Xinyue	Canada	$15,078	32
	Xiaowei	England	$4,509	21
	Xiao	Canada	$93,291	44
	Xiao	USA	$80,914	32
	Wenyue	Canada	$67,755	29
	Tatsuya	England	$51,324	28
	Syeed	USA	$74,144	35
	Peikun	Canada	$41,322	34
	Michaila	Canada	$36,409	24
	Larissa	Ireland	$89,600	27
	Kevin	USA	$61,515	43
	Junhyeok	Cnina	$65,624	28
	Joel	Canada	$65,257	44
	Jielin	England	$82,360	35
	Jiangyue	USA	$88,412	41
	Jia	Canada	$93,943	22
	Jason	Cnina	$57,397	43
	Jaclyn	Ireland	$89,043	37
	Irina	USA	$32,824	25

5.2. Sorting

MS 365 and Excel 2021 offer two powerful sorting operations, SORT and SORTBY, that use dynamic arrays. The advantage of these tools is that, when combined with naming an array, both functions update as the latest information enters the table (e.g., a change in a value or the addition of new data).

> See: Example Names, SORT, and SORTBY.xlsx

> Video: SORT and SORTBY

5.3. Frequency function

The final array function supports the creation of histograms without using the dialogue box under Data Analysis. This becomes useful in creating policy models where you need to analyse how changing assumptions affect model results. The video and spreadsheet also illustrate how to generate numbers for quick analysis and name ranges and vectors.

> See: Example Frequency Function.xlsx

> Video: The Frequency Function

Module 12: Matrices and Array Formulas – Applications in economics

6. Summary

This Module introduces Excel as a process for calculating higher order economic analytics using vectors and arrays. It also introduces array formulas and dynamic arrays. By design, it does not cover Pivot Tables, which offer powerful summarizing techniques without tedious formula writing. Every economist should have familiarity with Pivot Tables, but these are data management procedures and strictly do not involve economic analytics.

Annex: Key Excel formulas and functions

This Module introduced new functions, notably array formulas that make intelligent guesses about actions. Transforming Excel arrays into tables (assuming each column has a header [title]) and then naming the table offers important "resilience" since updates to the table (changing a value or adding information in the form of new rows or columns) will update all your calculations without the need for changing formulas in detail.

Function	Example	Explanation
=XLOOKUP(lookup-value, lookup-array, return-array,[if_not_found], [match_mode], [search_mode])	=XLOOKUP(B4,A7:A17,C7:C13,,-1)	This function searches two columns for the value in B4. The next two arguments are not used (for this example), and the search mode is for an exact match or the next smaller value.
=MMULT (array1, array2)	=MMULT(A1:C10:F3:O3)	This function multiplies two conformal matrices. Verify that the example has conformal matrices.
=SUMPRODUCT(array1,array2, array3…)	=SUMPRODUCT(A1:A10,B1:B10)	Multiply the vector A1:A10 by the vector A1:B10. Study the examples in the help area (F1) to see how powerful this formula can become.
=MINVERSE(array1)	=MINVERSE(C2:E4)	Take the inverse of matrix (array). This assumes a determinant exists and that array1 is square.

Module 12: Matrices and Array Formulas – Applications in economics

Function	Example	Explanation
=MDETERM(array1)	=MDETER(F4:I7)	This function computes the determinant of a square matrix. If rows or columns are linearly dependent, the determinant does not exist and a "VALUE!" error results.
=TRANSPOSE(array1)	=TRANSPOSE(A1:U300)	This takes the transpose of an array, which need not be square.
=FREQUENCY(data array, bin array)	=FREQUENCY(A1:A3003,J1:J10)	Developing a frequency distribution and histogram becomes easy with this function. This function is useful in probabilistic models.
Note: 1) Array is always specific to the function; 2) Naming arrays will clarify your code; and 3) Making arrays into Excel tables, labelling columns (or rows) can increase flexibility and leads to using Pivot tables for reporting.		