Linear Algebra

B A Ødegaard

24 November 2021

1 Introduction, why do we need to know linear algebra?

I prefer to think of Linear Algebra as a very convenient shorthand for mathematical operations on data.

Traditionally, much of introductory courses in linear algebra were concerned with the actual mechanics of calculation of some simple properties of matrices, but may never have given "the big picture" in terms of understanding linear algebra by analogy with the real numbers.

These days, the impetus to know and use linear algebra comes from convenience when one want to do nontrivial calculations on a computer. A tool implementing numerical linear algebra operations in a compact way has turned out to be very convenient for working on problems in finance and economics.

It is also very hard to do any kind of econometrics without knowing the basics of linear algebra. I will also claim that it is almost impossible to *understand* the intuition behind the results without a basic understanding of linear algebra.

I first go over some of the basic results of linear algebra. I am not going to go into proofs, but the results should be known. I will first go over the theoretical results at high speed, and then slow down, go back, and illustrate these results using actual numbers and the computer for calculations.

2 Theoretical summary of Linear Algebra

I now summarize the theoretical concepts of linear algebra. My main goal here is to illustrate the similarities between the operations in linear algebra and the standard operations that are used on real numbers.

While I do not want to get to formal, it will not hurt to see the abstract definitions of a number of concepts.[1] They are put here mainly for reference.

2.1 Fields

A set of elements $e_1, e_2, \cdots$ are said to belong to a *field* (F), if they are closed under the operations of addition $(e_j + e_i)$ and multiplication, $(e_i e_j)$, that is, the sum and products of elements of F also belong to F, and satisfy the following conditions. (A for addition, M for multiplication).

$$\begin{array}{lll}
(A1) & e_i + e_j = e_j + e_i & \text{(commutative law)} \\
(A2) & e_i + (e_j + e_k) = (e_i + e_j) + e_k & \text{(associative law)} \\
(A3) & \text{For any two elements } e_i,\ e_j, & \\
 & \text{there exists an element } e_k & \\
 & \text{such that } e_i + e_k = e_j &
\end{array}$$

[1] These definitions are from Rao (1973).

This condition implies that there is an element e_0 such that $e_i + e_0 = e_i$ for all i. The element e_0 is like zero of the (real) number system.

$(M1)$	$e_i e_j = e_j e_i$	(commutative law)
$(M2)$	$e_i(e_j e_k) = (e_i e_j)e_k$	(Associative law)
$(M3)$	$e_i(e_j + e_k) = e_i e_j + e_i e_k$	(Distributive law)
$(M4)$	For any two elements e_j and e_j such that $e_i \neq e_0$, the zero element there exists an element e_k such that $e_i e_k = e_j$	

$(M4)$ implies that there is an element e_1 such that $e_i e_1 = e_i$ for all i. This element e_1 is like 1 (unity) of the number system.

Example

The set of real numbers, denoted by $\mathcal{R}$, is the best known example of a *field*.

$$2 + 0 = 2$$

$$2 + 3 = 5 = 3 + 2$$

$$2 \cdot 1 = 2$$

$$2 \cdot 3 = 6 = 3 \cdot 2$$

2.2 Linear Vector Spaces.

For each of these definitions, see if you can see the connection to what you know about *real numbers*.

Consider a collection of elements $X, Y, Z, \ldots$, finite or infinite, which we choose to call vectors.[2]. Assume a set of scalar numbers $c_1, c_2, \ldots$, that constitute a *field* (Like the ordinary numbers with the operations of addition, subtraction, multiplication and division suitably defined). To define a linear space we lay down certain rules for combining the elements.

Vector Addition. The operation of addition indicated by $+$ is defined for any two vectors leading to a vector in the set and is subject to the following rules.

(i)	$X + Y = Y + X$	Commutative Law
(ii)	$X + (Y + Z) = (X + Y) + Z$	Associative Law

Null Element. There exists an element in the set, denoted by $\mathbf{0}$, such that

$$(iii) \qquad X + \mathbf{0} = X, \qquad \text{for all } X$$

Inverse (negative) Element. For any given element X, there exist a corresponding element ψ such that

$$(iv) \qquad X + \psi = \mathbf{0}$$

Scalar Multiplication. The multiplication of a vector X by a scalar c leads to a vector in the set, represented by cX, and is subject to the following rules.

(v)	$c(X + Y) = cX + cY$	(Distributive law for vectors)
(vi)	$(c_1 + c_2)X = c_1 X + c_2 X$	(Distributive law for scalars)
(vii)	$(c_1(c_2 X) = (c_1 c_2)X$	(Associative law)
$(viii)$	$1X = X$	1 is the scalar unit element.

[2]Note that this definition is not limited to what we usually think of as *vectors*, a set of n real numbers, $X = (x_1, x_2, \cdots, x_n)$. (The Euclidean Linear Space). It allows for more abstract elements, but you may want to think of the Euclidean case for intuition

A collection of elements (with the associated field of scalars F) satisfying axioms (i) to $(viii)$ above is called a *linear vector space*.

Some further useful definitions.

Dependence. A set of vectors $X_1, \cdots, X_n$ is said to be linearly dependent if there exists scalars $c_1, \cdots, c_n$, not all zero, such that

$$c_1 X_n + c_2 X_2 + \cdots + c_n X_n = 0$$

Otherwise it is independent.

Basis. A linearly independent subset of vectors in a vector space, generating or spanning the vector space, is called a *basis*.

Inner product. The inner product, denoted $(\cdot, \cdot)$, is a (complex or real-valued) function of two vectors, satisfying the following conditions

$$\begin{array}{ll} (i) & (X,Y) = \overline{(Y,X)}, \text{ the complex conjugate of } (X,Y) \\ (ii) & (X,X) > 0 \text{ if } X \neq \mathbf{0}, = 0 \text{ if } X = \mathbf{0} \\ (iii) & (cX,Y) = c(X,Y) \quad (\ c \text{ is a scalar }) \\ (iv) & (X+Y,Z) = (X,Z) + (Y,Z) \end{array}$$

Norm. The square root of (X,X) is called the *norm* of X and denoted by $\parallel X \parallel = \sqrt{(X,X)}$. The function $\parallel X - Y \parallel$ satisfies the postulates of a distance in a metric space.

Orthogonality. Two vectors X and Y are said to be *orthogonal* if their inner product, $(X,Y) = 0$.

Angle. The angle θ, between to non-null vectors X and Y is denoted by

$$\cos\theta = \frac{(X,Y)}{\parallel X \parallel \cdot \parallel Y \parallel}$$

Orthogonal basis. A set of non-null vectors $X_1, \cdots, X_n$, orthogonal in pairs, is necessarily independent. A set of orthogonal vectors spanning a vector space is called an orthogonal basis of the vector space.

Example

Let us look at the usual (Euclidean) vector space

$$\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$\mathbf{y} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Addition:

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2+1 \\ 1+2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 1+2 \\ 2+1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \mathbf{y} + \mathbf{x}$$

Null element

$$\mathbf{x} + \mathbf{0} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \mathbf{x}$$

Additive inverse element:

$$\mathbf{x} + \begin{pmatrix} -2 \\ -1 \end{pmatrix} = \mathbf{0}$$

Scalar multiplication

$$2\mathbf{x} = \begin{pmatrix} 2 \cdot 2 \\ 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

The two vectors

$$\begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

are dependent, because

$$1\begin{pmatrix}2\\2\end{pmatrix}+(-2)\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}=\mathbf{0}$$

Inner product (dot product)

$$\begin{pmatrix}2\\2\end{pmatrix}\cdot\begin{pmatrix}1\\1\end{pmatrix}=2\cdot 1+2\cdot 1=2+2=4$$

Norm

$$\left\|\begin{pmatrix}2\\2\end{pmatrix}\right\|=\sqrt{\begin{pmatrix}2\\2\end{pmatrix}\cdot\begin{pmatrix}2\\2\end{pmatrix}}=\sqrt{2\cdot 2+2\cdot 2}=\sqrt{4+4}=\sqrt{8}=2.8284$$

Orthogonal basis.

For the 2-dimensional case, we can generate all vectors from the two vectors

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

Why?

These are orthogonal, because

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

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

is thus an orthogonal basis.

2.3 Theory of Matrices

A **matrix A** is a collection of pq elements arranged in p rows and q columns

$$\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1q}\\ a_{21} & a_{22} & \cdots & \\ \vdots & & \ddots & \\ a_{p1} & \cdots & & a_{pq}\end{bmatrix}$$

We will usually write a matrix by bold capitals ($\mathbf{A}$). If necessary, will write explicitly the matrix as the collection of elements $[a_{ij}]$ where i indexes rows, and j columns.

The following operations are defined on matrices.

Addition. If $\mathbf{A}$ and $\mathbf{B}$ are two matrices with the same dimension $p\times q$, then the sum of $\mathbf{A}$ and $\mathbf{B}$ is defined to be the matrix $\mathbf{A}+\mathbf{B}=[a_{ij}+b_{ij}]$, of the same order $p\times q$. You may want to verify that

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$$

$$\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+\mathbf{B}+\mathbf{C}$$

Example

$$\mathbf{A}=\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}$$

$$\mathbf{B}=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$$

Addition

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix}2 & 3\\ 3 & 5\end{bmatrix}=\mathbf{B}+\mathbf{A}$$

Scalar multiplication. As for vectors, a matrix may be multiplied by a scalar c, the result being a matrix of the same form

$$c\mathbf{A} = [c \cdot a_{ij}]$$

Given this definition, it is easy to verify that

$$(c+d)\mathbf{A} = c\mathbf{A} + d\mathbf{A}$$

$$c(\mathbf{A}+\mathbf{B}) = c\mathbf{A} + c\mathbf{B}$$

Example

$$2\mathbf{A} = \begin{bmatrix} 2\cdot 1 & 2\cdot 2 \\ 2\cdot 2 & 2\cdot 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}$$

Matrix Multiplication. The matrix product $\mathbf{AB}$ is defined only when the number of columns in $\mathbf{A}$ is the same as that of the rows in $\mathbf{B}$. The result is a matrix with the (i, j) entry as the (inner) product of the i'th row vector of $\mathbf{A}$ with the j'th column vector of $\mathbf{B}$.

If $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$, then the product is obtained as

$$\mathbf{AB} = [c_{ij}] \quad c_{ij} = \sum_r a_{ir} b_{rj}$$

If $\mathbf{A}$ is $p \times q$ and $\mathbf{B}$ is $q \times t$, then $\mathbf{C} = \mathbf{AB}$ is $p \times q$.

Example

$$\mathbf{AB} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1\cdot 1 + 2\cdot 1 & 1\cdot 1 + 2\cdot 1 \\ 2\cdot 1 + 4\cdot 1 & 2\cdot 1 + 4\cdot 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 6 & 6 \end{bmatrix}$$

Example

As a more involved example, consider

$$\begin{bmatrix} -1 & 2 & 3 \\ 5 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 4 & 1 & 3 \\ 3 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 16 & 3 & 12 \\ 15 & 13 & 7 \end{bmatrix}$$

Note that *the commutative law does not hold for matrix multiplication.* In general $\mathbf{AB} \neq \mathbf{BA}$. It may also be that $\mathbf{AB}$ is defined but not $\mathbf{BA}$. The associative law is however satisfied,

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} = \mathbf{ABC}$$

In the cases where the matrices are compatible, the distributive law also holds,

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

$$(\mathbf{A}+\mathbf{B})(\mathbf{C}+\mathbf{D}) = \mathbf{A}(\mathbf{C}+\mathbf{D}) + \mathbf{B}(\mathbf{C}+\mathbf{D})$$

Null matrix. A null matrix, denoted by $\mathbf{0}$, is a matrix with all elements equal to zero.

$$\mathbf{0} = \begin{bmatrix} 0 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & 0 \end{bmatrix}$$

If $\mathbf{0}$ is the same dimension as $\mathbf{A}$

$$\mathbf{A} + \mathbf{0} = \mathbf{A}$$

Identity matrix. The *identity matrix*, denoted by $\mathbf{I}$, is a square matrix with ones along the diagonal, and zeros elsewhere. If we want to be explicit about the dimension, we will use the symbol $\mathbf{I}_q$ of the identity matrix with dimension $q \times q$.

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \\ 0 & & & & 0 & 1 \end{bmatrix}$$

If the operations are defined

$$\mathbf{IA} = \mathbf{A}$$

$$\mathbf{AI} = \mathbf{A}$$

Thus the matrices $\mathbf{0}$ and $\mathbf{I}$ behaves like 0 and 1 in the (usual) number system, but remember that the operations may not always be defined.

Transpose of a matrix. The matrix $\mathbf{A}'$ (also denoted $\mathbf{A}^T$), obtained by interchanging the rows and columns of $\mathbf{A}$ is called the *transpose* of $\mathbf{A}$. This means that the (i, j)'th entry of $\mathbf{A}$ is the same as the (j, i)'th entry of $\mathbf{A}'$. A useful property of transposes is that

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

Note: if $\mathbf{X}$ and $\mathbf{Y}$ are column vectors, the inner product of these vectors is equal to

$$\mathbf{X}'\mathbf{Y} = \mathbf{Y}'\mathbf{X}$$

Also, if $\mathbf{A}$ is a matrix, and $\mathbf{X}$ a column vector, then the product $\mathbf{AX}$ is a linear combination of the columns of $\mathbf{A}$.

Rank of a matrix. A matrix may be considered a collection of (row or column) vectors. The *rank* of the matrix is the number of linearly independent rows in the matrix, which is equal to the number of independent columns in the matrix.

Example

$$\text{rank} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = 2$$

$$\text{rank} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = 1$$

Inverse of a matrix. A square matrix of order $(m \times m)$ is said to be nonsingular if its rank is m. Otherwise it said to be singular. If a matrix $\mathbf{A}$ is nonsingular, there exists a unique matrix $\mathbf{A}^{-1}$, called the *inverse* of $\mathbf{A}$, such that

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_m$$

A useful result about inverses is that

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

Example

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\mathbf{A}^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

Let us check this

$$\mathbf{A} \cdot (\mathbf{A}^{-1}) = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + 1 \cdot (-1) & 1 \cdot (-1) + 1 \cdot 1 \\ 1 \cdot 2 + 1 \cdot (-1) & 1 \cdot (-1) + 2 \cdot 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Generalized inverse of a matrix. The inverse is only defined for square nonsingular matrices. A much more general concept is the g-inverse. If $\mathbf{A}$ is a $m \times n$ matrix of any rank, there exist a matrix $\mathbf{A}^{-}$ such that $\mathbf{X} = \mathbf{A}^{-}\mathbf{Y}$ is a solution of the equation $\mathbf{AX} = \mathbf{Y}$ for any $\mathbf{Y}$ in the span of $\mathbf{A}$. Note that the g-inverse defined this way is not necessarily unique.

Idempotent matrix. A matrix that multiplied with itself does not change we call *idempotent.* $\mathbf{A}$ is idempotent if

$$\mathbf{AA} = \mathbf{A}$$

The **Determinant** of a square matrix $\mathbf{A}$, of order m, written $|A|$, is a real valued function of the elements a_{ij} defined by:

$$|\mathbf{A}| = \sum \pm a_{1i} a_{2j} \cdots a_{mp}$$

where the summation is taken over all permutations $(i, j, \cdots, p)$ of $(1, 2, \cdots, m)$, and the sign is $+$ if the permutation is odd, $-$ if it is even.

Some properties of determinants:

- $|\mathbf{A}| = 0$ only if $\text{rank}(\mathbf{A}) < m$. (Hence if $|\mathbf{A}| \neq 0$, then $\mathbf{A}$ is invertible.)
- if c is a constant,

$$|c\mathbf{A}| = |c|^m \cdot |\mathbf{A}|$$

- if $\mathbf{A}$ is diagonal or triangular, $|\mathbf{A}|$ is the product of the diagonal elements.

Example

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$|\mathbf{A}| = (1 \cdot 2 - 1 \cdot 1) = 1$$

Square Root of Matrix. For a matrix $\mathbf{A}$, if there exist another matrix $\mathbf{A}^{\frac{1}{2}}$ such that

$$\mathbf{A}^{\frac{1}{2}} \cdot \mathbf{A}^{\frac{1}{2}} = \mathbf{A}$$

we call this the "square root" matrix.

2.4 Systems of Linear Equations.

Let $A_1, A_2, \cdots A_n$ be vectors. Consider the linear equation in the scalars $x_1, x_2 \cdots x_n$

$$A_1 x_1 + A_2 x_2 + \cdots A_n x_n = \mathbf{0} \tag{1}$$

A necessary and sufficient condition that this has a nontrivial solution (that is, not all equal to zero) is that $A_1 \cdots A_n$ be dependent.

Claim: *If there is more than one solution to a set of linear equations, then there is an infinite number of solutions.*

Proof: Let $X = \{x_1, \cdots x_n\}$ and $Y = \{y_1, \cdots, y_n\}$ solve the linear equation, $X \neq Y$. Then

$$A_1 x_1 + A_2 x_2 + \cdots + A_n x_n = \mathbf{0}$$

$$A_1 y_1 + A_2 y_2 + \cdots + A_n y_n = \mathbf{0}$$

Since both equations equal zero, we can multiply each equation with a scalar without changing the sum. Let λ be some real number.

$$\lambda \left(A_1 x_1 + A_2 x_2 + \cdots A_n x_n\right) = \mathbf{0}$$

$$(1 - \lambda) \left(A_1 y_1 + A_2 y_2 + \cdots + A_n y_n\right) = \mathbf{0}$$

Add the two equations:

$$A_1(\lambda x_1 + (1-\lambda)y_1) + \cdots + A_n(\lambda x_n + (1-\lambda)y_n) = \mathbf{0}$$

Since this hold for any real λ, there is an infinite number of solutions to the set of linear equations.

Consider next the *non-homogeneous* equation

$$A_1x_1 + A_2x_2 + \cdots A_nx_n = A_{n+1} \tag{2}$$

This equation 2 admits a solution if and only if A_{n+1} is dependent on $A_1 \cdots A_n$.

A system of equations is most easily written using matrices. If the vectors are the usual (Euclidean) case, the system is written as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \vdots &\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Define the matrices

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

The system of equations is then compactly written as

$$\mathbf{A}X = 0$$

Similarly, to write

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n &= b_2 \\ \vdots &\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n &= b_n \end{aligned}$$

in matrix notation, we define the column vector

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

and define the system of equations as

$$\mathbf{A}X = B$$

The results above about solvability is then expressed as conditions on the ranks of the matrices.

The homogeneous equation:

$$\mathbf{A}X = \mathbf{0}$$

$\mathbf{A}$ is an $n \times m$ matrix of rank $r \leq \min(n, m)$. The set X of all solutions is a linear subspace of rank $m - r$.

If $\mathbf{A}$ is square, and the rank r equals the number of equations, the equation only admits the trivial solution $X = \mathbf{0}$.

The non-homogeneous equation:

$$\mathbf{A}X = B$$

only admits a solution the rank of the matrix $\mathbf{A}$ is equal to the rank of the matrix defined by augmenting $\mathbf{A}$ with the column vector B.

The simplest case is when $\mathbf{A}$ is a square matrix, say $n \times n$. Then the solution of $\mathbf{A}X = B$ has a unique solution if $\text{rank}(\mathbf{A}) = n$, (the matrix is non-singular) and the solution is

$$X = \mathbf{A}^{-1}B$$

3 Using the computer for linear algebra, Matrix tools

A large number of computing environmenents now is available for dealing with linear algebra on actual numbers. Such an environment is the next logical step when spreadsheets become too limited, as you can at low cost deal with complicated calculations. Most of these tools tend to look very much alike, as long as you can get data into matrices you will manupulate the matrices using commands. In this chapter we will use one particular program to illustrate with, `octave`. This has the benefit of being freely available. Similar freely available tools are `scilab` and R. Commercial offerings include `matlab` and `S`.

4 Matrix theory using a computer tool

We will now go over these basic matrix operations and results with actual computer examples.

4.1 Starting the program

This is left as an exercise to the reader. It depends on what operating environment you are working on. For users on a Windows type system you will find it on the menu, if properly installed. If on the other hand you are on a UNIX workstation with X windows installed you will then start `octave` from an `xterm` window. Type `octave` at the prompt, and the program starts, you get the prompt, which we will show here as

```
>>
```

which means that the program is ready to receive commands. .

In the text output of octave will be shown typewritten as:

```
>>  A = [1, 2, 3
>   4, 5, 6]
```

This particular command defines a matrix `A`, and `octave` responds with printing the matrix that was just defined:

```
A =
  1 2 3
  4 5 6
```

4.2 Basic matrix operations

4.2.1 Defining variables

The basic variables in `octave` are numbers, vectors and matrices. The type of each variable is determined as you define it:

```
>> a=1
a = 1
>> b=2
b = 2
>> c=3
c = 3
>> y=[1;2;3]
y =
  1
  2
  3
>> x=[1,2,3]
```

```
x =
  1  2  3
>> A=[1 2 3;4 5 6]
A =
  1  2  3
  4  5  6
```

As you see, a space or a comma means a new number, a semicolon a new row. To suppress the printing of what you just defined, end the line with a semicolon:

```
>> A=[1,2,3,4];
>> A=[1,2,3,4]
A =
  1  2  3  4
>>
```

You can also use defined variables to define new variables, as long as the dimensions make sense. For example, given the above definitions:

```
>>B=[c x]
B =
  3  1  2  3
>> C = [A;x]
C =
  1  2  3
  4  5  6
  1  2  3
>> D = [A y]
error: number of rows must match
error: evaluating assignment expression near line 22, column 3
```

If the dimensioning is wrong, you get an error message, and the variable is not defined.

To see what is in a variable, tell `octave` to print the value by giving the name:

```
>> a
a = 1
>> b
b = 2
>> A
A =
  1  2  3
  4  5  6
>> B
B =
  5  1  2  3
```

Note that `octave` is case-sensitive, both `A` and `a` are defined.

4.2.2 Arithmetic Matrix Operations

We now get to the important parts of `octave`, namely, its built–in matrix arithmetic.

Addition and subtraction.

Numbers

```
>> a=1;
>> b=2;
>> c=3;
>> a+b
ans = 3
>> b+c
ans = 5
```

Vectors

```
>> x=[1 2 3 4 ]
x =
  1  2  3  4
>> y=[4 3 2 1]
y =
  4  3  2  1
>> x+y
ans =
  5  5  5  5
>> y-x
ans =
   3   1  -1  -3
```

We can also take linear combinations of vectors:

```
>> a*x + b*y
ans =
  9  8  7  6
>> 0.5*x -1.5*y
ans =
  -5.50000  -3.50000  -1.50000   0.50000
```

Matrices

```
>> A=[1 2 3; 4 5 6];
>> B=[6 5 4; 3 2 1];
>> A+B
ans =
  7  7  7
  7  7  7
>> A-B
ans =
  -5  -3  -1
   1   3   5
>> a*A + b*B
ans =
  13  12  11
  10   9   8
```

4.2.3 Matrix Multiplication

When multiplying matrices, you need to be more careful about dimensions, but otherwise it is just like vector multiplication.

```
>> A=[1 2 3;4 5 6]
A =
  1  2  3
  4  5  6
>> B = [1 2;3 4; 5 6]
B =
  1  2
  3  4
  5  6
>> A*B
ans =
  22  28
  49  64
>> B*A
ans =
   9  12  15
  19  26  33
  29  40  51
```

For these matrices, both **AB** and **BA** are defined operations, but note that the results are different, in fact, even the dimension of the product is different.

If we let **B** be a 2×2 matrix, then multiplying **AB** is an error.

```
>> B=[1 2;3 4]
B =
  1  2
  3  4
>> A*B
error: nonconformant matrices (op1 is 2x3, op2 is 2x2)
>> B*A
ans =
   9  12  15
  19  26  33
```

4.2.4 Matrix Operations

Transpose

The transpose of a matrix `a` is `a'`:

```
>> A
A =
  1  2  3
  4  5  6
>> A'
ans =
  1  4
  2  5
  3  6
```

Null and identity matrices

```
>> null = zeros(3,3)
null =
  0  0  0
```

```
  0  0  0
  0  0  0
>> ident = eye(3,3)
ident =
  1  0  0
  0  1  0
  0  0  1
```

Rank:

```
>> A
A =
  1  2  3
  4  5  6
>> rank(A)
ans = 2
```

Inverse

```
>> D = [3 4;4 6]
D =
  3  4
  4  6
>> inverse(D)
ans =
   3.0000  -2.0000
  -2.0000   1.5000
```

To make sure that this *is* the inverse, multiply `D` and `inverse(D)`:

```
>> D * inverse(D)
ans =
  1  0
  0  1
```

Idempotent matrix

So that you can see that there really exists idempotent matrices, consider

$$\mathbf{F} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

```
>> F = [0.5 0.5; 0.5 0.5]
F =
  0.50000  0.50000
  0.50000  0.50000
>> F * F
ans =
  0.50000  0.50000
  0.50000  0.50000
```

Determinant

```
>> B
B =
  1  2
```

```
  3  4
>> det(B)
ans = -2
```

A Warning. Numerical errors.

One thing you should be aware of when using a tool like `octave` is that the underlying tools are floating point number routines. All these are subject to numerical round–off errors, that may aggregate, and give weird results. This is also a problem in doing numerical linear algebra, if a matrix is *ill-conditioned*. As an example, consider the matrix

```
>> A
A =
  1  2  3
  3  2  1
  1  1  1
```

This matrix is singular, which you see if you take $\frac{1}{4}$ times the sum of the first two rows, which is equal to the third row. Since the matrix is singular, it does not have an inverse, and should have determinant equal to zero. Let us check.

```
>> det(A)
ans =  -4.4409e-16
>> inverse(A)
ans =
  -2251799813685248  -2251799813685246  9007199254740989
  4503599627370496  4503599627370494  -18014398509481980
  -2251799813685248  -2251799813685248  9007199254740992
>> rank(A)
ans = 2
```

As you see, because of numerical inaccuracies, `octave` does not "catch" that the determinant of $\mathbf{A}$ is actually *equal to* zero, and tries to invert $\mathbf{A}$ giving this extremely wrong estimate of the inverse. The algorithm for finding the rank of the matrix, on the other hand, is accurate enough to find the correct rank of 2. Taking determinants is thus useful as a diagnostic tool.[3]

Solving linear equations.

Remember the basic linear equation

$$\mathbf{Ax} = \mathbf{b}$$

We saw that this equation had a defined solution if the rank of $\mathbf{A}$ was the same as the rank of $[\mathbf{A}|\mathbf{b}]$. If $\mathbf{A}$ is nonsingular, we solve the linear equation by finding the unique solution

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

[3]There are actually better tools available to check the condition of a matrix, which you may want to look at if you are doing important computations. These problems are not due to low quality of the matrix routines, which are actually the best US government money can buy. (The basic matrix operations are done using LAPACK, a set of efficient and accurate FORTRAN routines financed by the US government.)

Exercise 1.

Consider the linear equation

$$3x_1 + 4x_2 = 5$$

$$4x_1 + 6x_2 = 8$$

Solve this system of equations using a matrix tool.

Solution to Exercise 1.

Write this in matrix form by defining

$$\mathbf{A} = \begin{bmatrix} 3 & 4 \\ 4 & 6 \end{bmatrix}$$

$$b = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

Let us first check that this system is solvable

```
>> A = [3 4;4 6]
A =
  3  4
  4  6
>> b=[5;8]
b =

  5
  8
>> rank(A)
ans = 2
>> rank ([A b])
ans = 2
```

Note how I create the augmented matrix $[A|b]$ by `[A b]`. The rank of the two is the same. Since $\mathbf{A}$ is square, we can calculate the solution as

```
>> inverse(A)
ans =
   3.0000  -2.0000
  -2.0000   1.5000
>> x = inverse(A) * b
x =
  -1
   2
```

The solution to the system of equations is

$$\mathbf{x} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

(In this case we calculated the solution by finding the inverse. But you should be aware that solving the system of equations by calculation of the inverse is not the numerically most stable way of doing the calculation. `Matlab/Octave` has built in a direct linear matrix solver, which is invoked by the *left division* operator

```
>> x = A\b
x =
  -1
   2
```

This solves the system of equations directly, and it is usually the way to do this operation, unless one needs the inverse for other purposes.)

Readings. The basic definitions of vector spaces are taken from Rao (1973)

Theil (1971) 1.1, (1.2), 1.3, (1.4), (1.6) are basic sources on the use of linear algebra in econometrics.

Further Reading Gale (1960) is a good introduction, although somewhat dated for the rest of the material, to the use of linear algebra in economics. The same can be said of appendix B of Dorfman, Samuelson, and Solow (1987)

Any econometrics textbook will have an appendix with the basic linear algebra results, in more or less comprehensive form. The Davidson and MacKinnon (1993) appendix is useful, but terse.

Mathematical text, see Sydsæter and Øksendal (1977) or Hadley (1961).

Advanced Sources For a serious development of the mathematics of this material, the obvious first place to look is Rao (1973).

References

Russel Davidson and James G MacKinnon. *Estimation and Interference in Econometrics.* Oxford University Press, 1993.

Robert Dorfman, Paul A Samuelson, and Robert M Solow. *Linear Programming and Economic Analysis.* Dover, 1987. Originally published: MacGraw-Hill, 1958.

David Gale. *The Theory of Linear Economic Models.* University of Chicago Press, 1960.

G Hadley. *Linear Algebra.* Addison-Wesley, 1961.

C Radhakrisna Rao. *Linear Statistical Inference and its applications.* Wiley, Second edition, 1973.

Knut Sydsæter and Bernt Øksendal. *Lineær Algebra.* Universitetsforlaget, 1977.

Henri Theil. *Principles of econometrics.* Wiley, 1971.