M

THE UNIVERSITY OF MELBOURNE
FACULTY OF BUSINESS & ECONOMICS

Helpsheet

Giblin Eunson Library

MATRIX ALGEBRA

Use this sheet to help you:

- Understand the basic concepts and definitions of matrix algebra
- Express a set of linear equations in matrix notation
- Evaluate determinants
- Invert a 3 x 3 matrix
- Understand how matrix inversion may be used to solve a set of linear equations

Author: Carter, D.
Design and layout: Pesina, J.

library.unimelb.edu.au/libraries/bee

Definitions and notation

A matrix is an arrangement of numbers or symbols in rows and columns.

A 3 x 2 matrix

$$\begin{bmatrix} 1 & 6 \\ -4 & 9 \\ 11 & 17 \end{bmatrix}$$

A matrix with 3 rows and 2 columns

A 3 x 3 matrix called a **square matrix** because the number of rows = number of columns

$$\begin{bmatrix} 42 & 91 & 78 \\ 11 & -6 & 0 \\ -5 & 99 & 2 \end{bmatrix}$$

A matrix with 3 rows and 3 columns

The numbers 42, 91 and so on are called **elements** of the matrix.

A matrix, **A**, is of order m x n if it has m rows and n columns. It may help to remember this as rc. So the matrix:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

is of order 2 x 3 because it has 2 rows and 3 columns.

We normally use a bold capital letter to write the name of a matrix, such as A.

In the matrix, A, the elements $a_{11,}$ a_{12} and so on are identified by their subscripts. The first subscript gives the row, the second the column. Again it may help to remember this as rc. So in general a_{ij} is in the ith row and the jth column of the matrix, A.

Note that a matrix is uniquely defined by its elements, so if even one element changes then we have a new matrix. This means that if we have matrix A and another matrix B, then we can only write A = B if each element of A equals the corresponding element of B.

Two special cases of matrices

(1) The null matrix:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$ (every element is zero)

(2) The **unit**, or **identity**, matrix (which must be square):

(the 'leading' diagonal = 1, all others = 0)

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

Vectors

A matrix with only one row is called a **row** vector.
A matrix with only one column is called a **column** vector.

Row vector: **a** = $\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix}$

Column vector: **b** = $\begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$

We normally use a bold, lower case letter to denote a vector, such as a.

Rules for manipulation of matrices

Transpose of a matrix

When we transpose a matrix, the first row becomes the first column, the second row becomes the second column, and so on.

e.g. **A =** $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$

The transpose is **A' =** $\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}$

The transpose of A is also written as A^T.

Addition and subtraction of two matrices

The rule is that we add/subtract the elements one by one.
e.g. **A + B = C**

means: $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} =$

$$\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{bmatrix}$$

Addition/subtraction is impossible unless **A** and **B** are of the same order.

Multiplication of two matrices

This is easiest shown diagrammatically. Suppose we have a matrix **A** and a matrix **B**, and we want to multiply them together to create a new matrix **C**. Then the product:
A x B = C

is calculated as follows:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{21} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

The rule is that we construct c_{11}, the top left-hand element of the new matrix, as follows. We take the element in row 1 of **A** and multiply it by the corresponding element of column 1 of **B**. So a_{11} multiplies b_{11}, and a_{12} multiplies b_{21}. The results are then added and this becomes element c_{11} of the new matrix, **C**.

Thus $c_{11} = a_{11}b_{11} + a_{12}b_{21}$. Then to construct c_{12}, we again take each element in row 1 of **A** and multiply it by the corresponding element of column 2 of **B**. So a_{11} multiplies b_{12}, and a_{12} multiplies b_{22}. The results are then added and this becomes element c_{12} of the new matrix. Thus $c_{12} = a_{11}b_{12} + a_{12}b_{22}$. And so on.

Generalising, the sum of each of the elements of the ith row of **A** times each of the elements of the jth column of **B** gives the ijth element of C.
Therefore again we have the pattern of **rc**.

Pre- and post-multiplication

Above we calculated the product **A x B**. Notice that this is not the same as **B x A**. For applying the rule row by column, we find that:

B x A = D

Is calculated as:

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \times \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} =$$

$$\begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{21} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{bmatrix}$$

so we can see that the elements of **D** are not the same as the elements of **C**. So in general, in matrix multiplication **AB** is not equal to **BA**.

To distinguish the two cases, in **AB** we say that **B** is pre-multiplied by **A**, while in **BA** we say that **B** is post-multiplied by **A**.

Conformability

In order for the matrix product **AB** to exist we need the number of elements in the rows of **A** to be equal to the number of columns of **B**.

Thus we can say **AB** exists if the number of columns in **A** equals the number of rows in **B**. Generalising, if we are given two matrices, **A** of order m x s and **B** of order s x n, then we can say:

1) The product **AB** exists because the number of columns, s, in **A** is equal to the number of rows, s, in **B**, and also that the product **AB** will be a matrix of order m x n.

2) The product **BA** does not exist, unless m = n.

Vector multiplication

A special case of matrix multiplication is when **A** = 1 x n (a row vector) and **B** is n x 1 (a column vector) e.g if n = 3, then if we pre-multiply **B** by **A**, we get:

$$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$$

$$= a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}$$

Note that $a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}$ is not a matrix nor a vector. It is simply an ordinary number. In matrix algebra, an ordinary number is called a scalar. So in general a row vector multiplied by a column vector is a **scalar**.

In general a column vector multiplied by a row vector results in a matrix.

Scalar multiplication

This is when you multiply a matrix (A) by a constant or scalar (k).
The rule is simply that each element of **A** is multiplied by k. For example, if **A** is of order 2

$$k\mathbf{A} = k\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{bmatrix}$$

We can see here why k is called a scalar, as it scales up all the elements in A.

Matrix algebra as a compact notation

Suppose we have three simultaneous linear equations:
4x + y – 5z = 8
-2x + 3y + z = 12
3x – y + 4z = 5

If we define

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & -5 \\ -2 & 3 & 1 \\ 3 & -1 & 4 \end{bmatrix}, \ \boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \ \boldsymbol{b} = \begin{bmatrix} 8 \\ 12 \\ 5 \end{bmatrix}$$

Then we can write the simultaneous equations as :
Ax = b
Thus we see that matrix algebra gives a compact and easily manageable notation for handling large sets of linear simultaneous equations.

The determinant of a square matrix

Associated with every square matrix there is a single scalar called a determinant. How the determinant is defined and calculated will be shown using examples.

The determinant of a 2 x 2 matrix

Given: $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

then the determinant of A, written as det. A, or $|\mathbf{A}|$, is defined as: ad – bc.
If det. A = 0, the matrix A is said to be singular.

The determinant of a 3 x 3 matrix

Given: $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

Then we construct det A as follows:

Step 1: We take the first element of row 1, a_{11}. Then we delete the row and the column that include a_{11}: that is, the first row and the first column. This leaves us with a 2 x 2 matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Step 2: Then we find the determinant of this sub-matrix, which is $a_{22}a_{33}$ - a_{23} a_{32}. This determinant is called a minor, and is written $|\mathbf{M_{11}}|$, where the subscripts tell us that it is the determinant found after deleting row 1 and column 1.

Thus $|\mathbf{M_{11}}| = a_{22}a_{33} - a_{23}\,a_{32}$

Then we repeat steps 1 and 2, this time taking in step 1 the second element in row 1, a_{12}. In this case when we delete the row and column that include a_{12} the resulting sub-matrix is:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

so in step 2 the determinant of the sub-matrix is $a_{21}a_{33}$ - a_{23} a_{31} and the minor is $|\mathbf{M_{12}}| = a_{21}a_{33} - a_{23}a_{31}$.

Third we repeat steps 1 and 2 again, this time taking in step 1, the third element in row 1, a_{13}. In this case when we delete the row and column that include a_{13}, the resulting sub-matrix is:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

So in step 2 the determinant of the sub-matrix is $a_{21}a_{32}$ - a_{22} a_{31} and the minor is $|\mathbf{M_{13}}| = a_{21}a_{32} - a_{22}\,a_{31}$.

Finally we multiply each of the elements of row 1 by its associated minor, and add the results. Before doing this, however, we need to make one final adjustment. The minors are given a + sign if their subscripts add to an even number, and a – sign if their subscripts add to an odd number. So $|\mathbf{M_{11}}|$ and $|\mathbf{M_{13}}|$ retain their + signs (because 1 + 1 = 2 and 1 + 3 = 4 are both even numbers), while $|\mathbf{M_{12}}|$ is given a – sign because 1 + 2 = 3 is odd. The minors, after their signs have been adjusted in this way, are often renamed cofactors.

With these adjustments to signs, we now have the determinant of A as
$|\mathbf{A}| = \mathbf{a_{11}}|\mathbf{M_{11}}| - \mathbf{a_{12}}|\mathbf{M_{12}}| + \mathbf{a_{13}}|\mathbf{M_{13}}|$.

It is possible to find the determinant of higher order matrices in a similar manner, but is beyond what we need at this stage.

The inverse of a square matrix

Matrix $\mathbf{A^{-1}}$ is the inverse of matrix **A** if :
$\mathbf{A^{-1}A = I}$

where I is the unit or identity matrix. The process or technique for finding $\mathbf{A^{-1}}$ is called **matrix inversion.**

The following example is a demonstration of matrix inversion.

Previously we had a set of three simultaneous equations written as:

Ax = b, where

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & -5 \\ -2 & 3 & 1 \\ 3 & -1 & 4 \end{bmatrix}$$

To find the inverse, $\mathbf{A}^{-1}$, of **A**, we proceed as follows.

Step 1. Choose any row or column; let's say, row 1. Evaluate all the minors, $|\mathbf{M}_{11}|$, $|\mathbf{M}_{12}|$ and $|\mathbf{M}_{13}|$ of this row and arrange them as a matrix, with their appropriate signs according to whether the sum of the subscripts is even or odd. (These signed minors are called cofactors, and the matrix is called C, the matrix of cofactors.) The result is:

$$\mathbf{C} = \begin{bmatrix} \begin{vmatrix} 3 & 1 \\ -1 & 4 \end{vmatrix} & -\begin{vmatrix} -2 & 1 \\ 3 & 4 \end{vmatrix} & \begin{vmatrix} -2 & 3 \\ 3 & -1 \end{vmatrix} \\ -\begin{vmatrix} 1 & -5 \\ -1 & 4 \end{vmatrix} & \begin{vmatrix} 4 & -5 \\ 3 & 4 \end{vmatrix} & -\begin{vmatrix} 4 & 1 \\ 3 & -1 \end{vmatrix} \\ \begin{vmatrix} 1 & -5 \\ 3 & 1 \end{vmatrix} & -\begin{vmatrix} 4 & -5 \\ -2 & 1 \end{vmatrix} & \begin{vmatrix} 4 & 1 \\ -2 & 3 \end{vmatrix} \end{bmatrix}$$

Note the signs of the various elements: as explained earlier, each is positive or negative according to whether the sum of its subscripts is even or odd.

Step 2. Evaluate all the determinants, giving:

$$\mathbf{C} = \begin{bmatrix} 13 & 11 & -7 \\ 1 & 31 & 7 \\ 16 & 6 & 14 \end{bmatrix}$$

Step 3. Transpose C, to give C' (This is also called the adjoint matrix of A, written as adj. A.)

$$\mathbf{C}' = \begin{bmatrix} 13 & 1 & 16 \\ 11 & 31 & 6 \\ -7 & 7 & 14 \end{bmatrix}$$

Step 4. Evaluate $|\mathbf{A}|$. Using row 1 of A, we get:
$|\mathbf{A}| = 4(12 + 1) - 1(-8 - 3) + (-5)(2 - 9) = 98$

Step 5. Then A^{-1} is given by:

$$\mathbf{A}^{-1} = \frac{1}{|A|}\mathbf{C}' = \frac{1}{98}\begin{bmatrix} 13 & 1 & 16 \\ 11 & 31 & 6 \\ -7 & 7 & 14 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{13}{98} & \frac{1}{98} & \frac{16}{98} \\ \frac{11}{98} & \frac{31}{98} & \frac{6}{98} \\ \frac{-7}{98} & \frac{7}{98} & \frac{14}{98} \end{bmatrix}$$

We can check whether we have made any mistakes in calculating the inverse matrix $\mathbf{A}^{-1}$. If our calculations are correct, we should find that $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ (the unit matrix).

Use of matrix inversion to solve linear simultaneous equations

To demonstrate we will continue using the previous example.
We had three simultaneous equations

$4x + y - 5z = 8$
$-2x + 3y + z = 12$
$3x - y + 4z = 5$

which we could write as
Ax = b

where

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & -5 \\ -2 & 3 & 1 \\ 3 & -1 & 4 \end{bmatrix}, \; \boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \; \boldsymbol{b} = \begin{bmatrix} 8 \\ 12 \\ 5 \end{bmatrix}$$

If we find $\mathbf{A}^{-1}$ such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, we can pre-multiply both sides of equation (1) by $\mathbf{A}^{-1}$ and thereby get:

A^{-1}Ax = A-1b

But by definition $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, so substituting this on the left-hand side gives

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

But **Ix = x,** for any column vector **x**, so we have

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

which will give us the solution values of **x.**

So using the information from the previous example we have:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{13}{98} & \frac{1}{98} & \frac{16}{98} \\ \frac{11}{98} & \frac{31}{98} & \frac{6}{98} \\ \frac{-7}{98} & \frac{7}{98} & \frac{14}{98} \end{bmatrix} \begin{bmatrix} 8 \\ 12 \\ 5 \end{bmatrix}$$

Multiplying out the right-hand side.

$$x = \left(\frac{13}{98}\right)8 + \left(\frac{1}{98}\right)12 + \left(\frac{16}{98}\right)5 = 2$$

$$y = \left(\frac{11}{98}\right)8 + \left(\frac{31}{98}\right)12 + \left(\frac{6}{98}\right)5 = 5$$

$$z = \left(\frac{-7}{98}\right)8 + \left(\frac{7}{98}\right)12 + \left(\frac{14}{98}\right)5 = 1$$

By substituting these values into the original simultaneous equations we can easily check that these solutions for x, y and z are correct.