Matrix algebra

PART 2

[1]

[1] Please send comments, corrections and clarifications to peter.stewart@glasgow.ac.uk and gregory.stevenson@glasgow.ac.uk.

Section Overview

- **Key Points: Matrix algebra**
  - Definition of a matrix
  - Matrix addition and scalar multiplication
  - Matrix multiplication
  - Matrix powers
  - Transpose of a matrix
  - Inverse of a matrix
  - Determinant of a matrix
- **Associated sections of the book:**
  - Poole Section 3.1 (p138 - 152) (Omit partitioned matrices p145-149)
  - Poole Section 3.2 (p160) (Omit column operations)
  - Poole Section 3.3 (p169)

3.1: Matrix Operations

Definition A *matrix* is a rectangular array of numbers called the *entries* or *elements* of the matrix.

For a matrix A we write the i, j^{th} entry as a_{ij}.

For the matrix $A = \begin{bmatrix} 2 & 0 & 5 \\ 1 & 4 & -1 \end{bmatrix}$ then $a_{11} = 2$, $a_{23} = -1$, $a_{21} = 1$ and $a_{12} = 0$.

Symbolically this is expressed by $A = (a_{ij})$. So if A is an $m \times n$ matrix then

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

If the columns of A are the column vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ in $\mathbb{R}^m$ we write

$$A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n],$$

and if the rows are row vectors $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_m$ in $\mathbb{R}^n$ we write

$$A = \begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_m \end{bmatrix}.$$

A square matrix ($m = n$) is *diagonal* iff its off-diagonal elements are zero. The $n \times n$ identity matrix I_n is the diagonal matrix with all diagonal entries 1.

Example Consider the matrices $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 4 \\ 0 & 0 & 8 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Two matrices are *equal* if and only if they have the same size and their corresponding entries are equal.

Matrix addition and scalar multiplication

If $A = (a_{ij})$ and $B = (b_{ij})$ are $m \times n$ matrices and c is a scalar [2] and c is a scalar, then we can define new $m \times n$ matrices $A + B$ and cA componentwise: [3]

$$\begin{aligned} A + B &= (a_{ij} + b_{ij}), \\ cA &= c(a_{ij}) = (ca_{ij}). \end{aligned}$$

Matrix multiplication

If A is an $m \times n$ matrix and B is an $n \times r$ matrix then $C = AB$ is the $m \times r$ matrix with $(i, j)^{th}$ entry given by[4]

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Theorem 3.1 Let A be an $m \times n$ matrix, and

$$\mathbf{e}_i = [0, \cdots, 0, 1, 0, \cdots, 0],$$

[2] Note that matrices must be the same size for addition of two matrices to be defined.

[3] Note that we have the same two operations: addition and scalar multiplication defined for $m \times n$ matrices as we had in $\mathbb{R}^n$. These operations satisfy the same rules as our vectors did in Theorem 1.1 (see Theorem 3.2 below). So, the collection of $m \times n$ matrices is a *vector space* in the sense that we will define in chapter 6. We're starting to see the potential value in generalisation: the additive structure of matrices and of vectors is the same: we can (and will) prove general theorems which will simultaneously cover both situations.

[4] Again note that matrix multiplication is not defined for every pair (A, B) of matrices. It is only defined when the number of rows of A is equal to the number of columns of B. Mathematics is heavily typed: operations can only be performed when objects have the right type (which of course depends on the operation).

with the 1 in the i^{th} position, and

$$\mathbf{e}_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

with the 1 in the j^{th} position. Then

(a) $\mathbf{e}_i A$ is the i^{th} row of A,

(b) $A\mathbf{e}_j$ is the j^{th} column of A .

Proof: Omitted.

Matrix Powers

If A is an $n \times n$ matrix and k is a positive integer then

$$A^0 = \mathbb{I}_n, \qquad A^2 = AA, \qquad A^k = \underbrace{AA \cdots A}_{k \text{ times}}.$$

Example If $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ then compute A^2 and A^3.

The transpose of a matrix

Definition The *transpose* of an $m \times n$ matrix A is the $n \times m$ matrix A^{T} obtained by interchanging the rows and columns of A.

<u>Example</u> Compute the transpose of the matrices $A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$.

Definition A square matrix is *symmetric* if $A^T = A$. That is, A is equal to its own transpose.

<u>Example</u> Are the matrices $A = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 7 \\ 1 & 7 & 5 \end{bmatrix}$ symmetric?

3.2: Matrix Algebra

The addition and scalar multiplication rules for $m \times n$ matrices obey the following algebraic properties.[5]

[5] Note the similarity between this and Theorem 1.1.

Theorem 3.2 Let A, B and C be $m \times n$ matricies and c and d be scalars. Then

(a) $A + B = B + A$,

(b) $(A + B) + C = A + (B + C)$,

(c) $A + 0 = A$,

(d) $A + (-A) = 0$,

(e) $c(A + B) = cA + cB$,

(f) $(c + d)A = cA + dA$,

(g) $c(dA) = (cd)A$,

(h) $1A = A$.

Proof:

Properties of matrix multiplication

Matrix multiplication behaves differently from multiplication of numbers. In general multiplication is not commutative.[6] Also, we could have $A^2 = 0$ even if $A \neq 0$.[7]

[6] i.e. even for square $n \times n$ matrices A and B (so that AB and BA are defined and are matrices of the same size), it does not follow that AB and BA are equal.

[7] Can you give an example where this happens?

Theorem 3.3 Let A, B and C be matrices and k be a scalar. The following identities hold whenever the operations involved can be performed.

(a) $A(BC) = (AB)C$,

(b) $A(B + C) = AB + AC$,

(c) $(A + B)C = AC + BC$,

(d) $k(AB) = (kA)B = A(kB)$,

(e) $I_m A = A = AI_n \qquad \text{if } A \text{ is } m \times n$.

Proof: Omitted.

Similarly we have algebraic rules for the transpose.

Theorem 3.4 Let A and B be matrices. The following identities hold whenever the operations involved can be performed.

(a) $(A^{\mathrm{T}})^{\mathrm{T}} = A$,

(b) $(A + B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$,

(c) $(kA)^{\mathrm{T}} = k(A^{\mathrm{T}})$,

(d) $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$,

(e) $(A^m)^{\mathrm{T}} = (A^{\mathrm{T}})^m$ for all integers $m \geq 0$.

Proof: Omitted.

Theorem 3.5

(a) If A is a square matrix then $A + A^{\mathrm{T}}$ is a symmetric matrix,

(b) For any matrix A, AA^{T} and $A^{\mathrm{T}}A$ are symmetric matrices.

Proof: Omitted.

3.3: The inverse of a matrix

Definition If A is an $n \times n$ matrix, the *inverse* of A is an $n \times n$ matrix A' such that

$$AA' = I_n, \qquad \text{and} \qquad A'A = I_n.$$

If A' exists we say A is *invertible*. If no inverse exists, then we say that A is not invertible.

Example Consider the matrices $A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Theorem 3.6 If an $n \times n$ matrix A is invertible then its inverse is unique.

Proof: Omitted.

Notation If A is invertible we write A^{-1} for its inverse. **Important Warning.** We are not allowed to write $\frac{1}{A}$ for the inverse of A. Matrices are not numbers, we have defined a notion of multiplication but not of division.

Theorem 3.7 If A is an invertible $n \times n$ matrix then the system of linear equations given by $A\mathbf{x} = \mathbf{b}$ has the unique solution given by $\mathbf{x} = A^{-1}\mathbf{b}$.

Proof: Omitted.

Theorem 3.8 If
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
then A is invertible if $ad - bc \neq 0$, in which case
$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$
If $ad - bc = 0$ then A is not invertible.

Proof: Omitted.

For general $n \times n$ matrices, there are algorithms (using row reduction) for finding the inverse, but not convenient general formula as we have in the 2×2 case.

Definition For a 2×2 matrix
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$
we call $ad - bc$ the *determinant*[8] of A, so that
$$\det(A) = ad - bc.$$

[8] See section 4.2, where we will discuss the determinant of general square matrices

Example Find the inverses if they exist of $A = \begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 5 \\ 2 & 10 \end{bmatrix}$.

Example Solve the system

$$x + 5y = 3, \qquad 2x + 4y = 1$$

using the inverse of the coefficient matrix.

Theorem 3.9

(a) If A is an invertible matrix, then A^{-1} is invertible and

$$(A^{-1})^{-1} = A.$$

(b) If A is an invertible matrix and $c \neq 0$ is a scalar then cA is invertible and

$$(cA)^{-1} = \frac{1}{c}A^{-1}$$

(c) If A and B are invertible matrices of the same size, then AB is invertible and

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

(d) If A is an invertible matrix, then A^{T} is invertible and

$$(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}.$$

(e) If A is invertible matrix then A^n is invertible for all integers $n \geq 0$ and

$$(A^n)^{-1} = (A^{-1})^n.$$

Proof: Omitted.

Definition If A is invertible and $n \geq 0$ an integer we define

$$A^{-n} = (A^{-1})^n = (A^n)^{-1}.$$