Chapter 1

Matrix Analysis

1.1 Introduction

1.2 Review of Matrix Algebra

1.2.1 Definition

(a) An array of real numbers $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$ is called $m \times n$ *matrix*. The a_{ij} is referred to as the i, j-th element and denotes the element in the i-th row and j-th column. If $m = n$ then A is called a *square matrix of order* n. If the matrix has one column or one row then it is called a *column vector* or a *row vector* respectively.

(b) In a square matrix A of order n the diagonal containing the elements $a_{11}, a_{22}, \ldots, a_{nn}$ is called the principal or *leading diagonal*. The sum of the elements in this diagonal is called the *trace of* A, that is $\text{trace}\, A = \sum_{i=1}^{n} a_{ii}$.

(c) A *diagonal matrix* is a square matrix that has its only non-zero elements along the leading diagonal. A special case of a diagonal matrix is the *unit* or *identity matrix* I for which $a_{11} = a_{22} = \cdots = a_{nn} = 1$.

(d) A *zero* or *null matrix* 0 is a matrix with every element zero.

(e) The *transposed matrix* A^T is the matrix A with rows and columns interchanged, its i, j-th element being a_{ji}.

(f) A square matrix A is called a *symmetric matrix* if $A^T = A$. It is called *skew symmetric* if $A^T = -A$.

1.2.2 Basic Operations on Matrices

Equality

The matrices A and B are equal, that is $A = B$, if they are of the same order $m \times n$ and $a_{ij} = b_{ij}, \quad 1 \le i \le m, \quad 1 \le j \le n$.

Mutiplication by A Scalar

If λ is a scalar then the matrix λA has elements λa_{ij}.

Addition

We can only add an $m \times n$ matrix A to another $m \times n$ matrix B and the elements of the sum $A + B$ are $a_{ij} + b_{ij}, \quad 1 \le i \le m, \quad 1 \le j \le n$.

Properties of Addition

(i) commutative law: $A + B = B + A$

(ii) associative law: $(A + B) + C = A + (B + C)$

(iii) distributive law: $\lambda(A + B) = \lambda A + \lambda B, \lambda$ scalar.

Matrix multiplication

If A is an $m \times p$ matrix and B a $p \times n$ matrix then we define the product $C = AB$ as the $m \times n$ matrix with elements

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n.$$

Properties of multiplication

(i) The commutative law is not satisfied in general; that is, in general $AB \neq BA$. Order matters and we distinguish between AB and BA by the terminology: pre-multiplication of B by A to form AB and post-multiplication of B by A to form BA.

(ii) Associative law: $A(BC) = (AB)C$ If λ is a scalar then $(\lambda A)B = A(\lambda B) = \lambda AB$

(iii) Distributive law over addition: $(A + B)C = AC + BC$
$A(B + C) = AB + AC$
Note the importance of maintaining order of multiplication.

(iv) If A is an $m \times n$ matrix and if I_m and I_n are the unit matrices of order m and n respectively then $I_m A = AI_n = A$.

Properties of the Transpose

If A^T is the transposed matrix of A then

(i) $(A + B)^T = A^T + B^T$

(ii) $(A^T)^T = A$

(iii) $(AB)^T = B^T A^T$

1.2.3 Determinants

The determinant of a square $n \times n$ matrix A is denoted by det A or $|A|$. If we take a determinant and delete row i and column j then the determinant remaining is called the minor M_{ij} of the i, j-th element. In general we can take any row i (or column) and evaluate an $n \times n$ determinant $|A|$ as $|A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}$.
A minor multiplied by the appropriate sign is called the *cofactor* A_{ij} of the i , j-th element so $A_{ij} = (-1)^{i+j} M_{ij}$ and thus $|A| = \sum_{j=1}^{n} a_{ij} A_{ij}$.

Some useful properties

(i) $|A^T| = |A|$

(ii) $|AB| = |A||B|$

(iii) A square matrix A is said to be non-singular if $|A| \neq 0$ and singular if $|A| = 0$.

1.2.4 Adjoint and inverse matrices

Adjoint matrix

The *adjoint* of a square matrix A is the transpose of the matrix of cofactors, so for a 3×3 matrix A

$$\operatorname{adj} A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}^T$$

Properties

(i) If A is non-singular then $|A| \neq 0$ and $A^{-1} = (\operatorname{adj} A)/|A|$.

(ii) If A is singular then $|A| = 0$ and A^{-1} does not exist.

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

Matlab Practice

Do matrix operations in *Matlab*.

1.2.5 Linear aligns

In this section we reiterate some definitive statements about the solution of the *system of simultaneous linear aligns*

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_1 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_1
\end{aligned}
$$

or, in matrix notation,

$$
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}
$$

that is,

$$Ax = b$$

where A is the matrix of coefficients and x is the vector of unknowns. If $b = 0$ the aligns are called *homogeneous*, while if $b \neq 0$ they are called *nonhomogeneous* (or inhomogeneous). Considering individual cases:

Case (i) If $b \neq 0$ and $|A| \neq 0$ then we have a unique solution $x = A^{-1}b$.

Case (ii) If $b = 0$ and $|A| \neq 0$ we have the trivial solution $x = 0$.

Case (iii) If $b \neq 0$ and $|A| = 0$ then we have two possibilities: either the aligns are inconsistent and we have no solution or we have infinitely many solutions.

Case (iv) If $b = 0$ and $|A| = 0$ then we have infinitely many solutions.

Case (iv) is one of the most important, since from it we can deduce the important result that the homogeneous align $Ax = 0$ has a non-trivial solution if and only if $|A| = 0$.

Matlab Practice

Solve system of simultaneous linear equations.

1.2.6 Rank of a Matrix

If A and $(A : b)$ have different rank then we have no solution to $Ax + b$. If the two matrices have the same rank then a solution exists, and furthermore the solution will contain a number of free parameters equal to $(n - \operatorname{rank} A)$.

1.3 Vector Spaces

1.3.1 Definition

A real vector space V is a set of objects called vectors together with rules for addition and multiplication by real numbers. For any three vectors a, b and c in V and any real numbers α and β the sum $a + b$ and the product α a also belong to V and satisfy the following axioms:

(a) $a + b = b + a$

(b) $a + (b + c) = (a + b) + c$

(c) there exists a zero vector 0 such that $a + 0 = a$

(d) or each a in V there is an element $-a$ in V such that $a + (-a) = 0$

(e) $\alpha(a + b) = \alpha a + \alpha b$

(f) $(\alpha + \beta)a = \alpha a + \beta a$

1.3.2 Linear independence

The idea of linear dependence is a general one for any vector space. The vector x is said to be *linearly dependent* on $x_1, x_2, \cdots, x_m$ if it can be written as $x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m$ for some scalars $\alpha_1, \cdots, \alpha_m$. The set of vectors $y_1, y_2, \cdots, y_m$ is said to be *linearly independent* if and only if $\beta_1 y_1 + \beta_2 y_2 + \cdots + \beta_m y_m = 0$.

Example. Show that $e_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$ and $e_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$ form a linearly independent set and describe $S(e_1, e_2)$ geometrically.

If we can find a set B of linearly independent vectors $x_1, x_2, \cdots, x_n$ in V such that $S(x_1, x_2, \cdots, x_n) = V$ then B is called a *basis* of the vector space V. Such a basis forms a crucial part of the theory, since every vector x in V can be written uniquely as $x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$. The definition of B implies that x must take this form. To establish uniqueness, let us assume that we can also write x as $x = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$. Then, on subtracting, $0 = (\alpha_1 - \beta_1)x_1 + \cdots + (\alpha_n - \beta_n)x_n$ and since $x_1, \cdots, x_n$ are linearly independent, the only solution is $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \cdots$; hence the two expressions for x are the same. It can also be shown that any other basis for V must also contain n vectors and that any $n + 1$ vectors must be linearly dependent. Such a vector space is said to have dimension n (or infinite dimension if no finite n can be found). In a three-dimensional *Euclidean space*

$$e_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad e_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, \quad e_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

form an obvious basis, and

$$d_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad d_2 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \quad d_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

is also a perfectly good basis.

1.3.3 Transformations between bases

Since any basis of a particular space contains the same number of vectors, we can look at transformations from one basis to another. We shall consider a three-dimensional space, but the results are equally valid in any number of dimensions. Let e_1, e_2, e_3 and e_1', e_2', e_3' be two bases of a space. From the definition of a basis, the vectors e_1', e_2' and e_3' can be written in terms of e_1, e_2 and e_3 as

$$\left.\begin{aligned} e_1' &= a_{11}e_1 + a_{21}e_2 + a_{31}e_3 \\ e_2' &= a_{12}e_1 + a_{22}e_2 + a_{32}e_3 \\ e_3' &= a_{13}e_1 + a_{23}e_2 + a_{33}e_3 \end{aligned}\right\}$$

Taking a typical vector x in V, which can be written both as

$$x = x_1 e_1 + x_2 e_2 + x_3 e_3$$

and as

$$x = x_1' e_1' + x_2' e_2' + x_3' e_3'$$

we can use the transformation to give

$$\begin{aligned} x &= x_1'(a_{11}e_1 + a_{21}e_2 + a_{31}e_3) + x_2'(a_{12}e_1 + a_{22}e_2 + a_{32}e_3) + x_3'(a_{13}e_1 + a_{23}e_2 + a_{33}e_3) \\ &= (x_1'a_{11} + x_2'a_{12} + x_3'a_{13})e_1 + (x_1'a_{21} + x_2'a_{22} + x_3'a_{23})e_2 + (x_1'a_{31} + x_2'a_{32} + x_3'a_{33})e_3 \end{aligned}$$

Then

$$\begin{aligned} x_1 &= a_{11}x_1' + a_{12}x_2' + a_{13}x_3' \\ x_2 &= a_{21}x_1' + a_{22}x_2' + a_{23}x_3' \\ x_3 &= a_{31}x_1' + a_{32}x_2' + a_{33}x_3' \end{aligned}$$

or $x = Ax'$.

Thus changing from one basis to another is equivalent to transforming the coordinates by multiplication by a matrix, and we thus have another interpretation of matrices. Successive transformations to a third basis will just give $x' = Bx''$, and hence the composite transformation is $x = (AB)x''$ and is obtained through the standard matrix rules.

For convenience of working it is usual to take mutually orthogonal vectors as a basis, so that $e_i^T e_j = \delta_{ij}$ and $e_i'^T e_j' = \delta_{ij}$, where δ_{ij} is the *Kronecker delta*

$$d_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

Multiplying out these orthogonality relations, we have

$$e_1'^T e_j' = \sum_k a_{ki} e_k^T \sum_p a_{pj} e_p = \sum_k \sum_p a_{ki} a_{pj} e_k^T e_p = \sum_k \sum_p a_{ki} a_{pj} \delta_{kp} = \sum_k a_{ki} a_{kj}$$

Hence

$$\sum_k a_{ki} a_{kj} = \delta_{ij}$$

or in matrix form

$$A^T A = I$$

It should be noted that such a matrix A with $A^{-1} = A^T$ is called an *orthogonal matrix*.

Exercises

1. To be written!

1.4 The Eigenvalue Problem

A problem that leads to a concept of crucial importance in many branches of mathematics and its applications is that of seeking non-trivial solutions $x \neq 0$ to the matrix equation

$$Ax = \lambda x$$

This is referred to as the eigenvalue problem; values of the scalar λ for which nontrivial solutions exist are called *eigenvalues* and the corresponding solutions $x \neq 0$ are called the *eigenvectors*.

1.4.1 The characteristic equation

The set of simultaneous equations

$$Ax = \lambda x$$

where A is an $n \times n$ matrix and $x = [x_1 x_2 \cdots x_n]^T$ is an $n \times 1$ column vector can be written in the form

$$(\lambda I - A)x = 0$$

where I is the identity matrix. We know that a non-trivial solution exists if

$$c(\lambda) = |\lambda I - A| = 0$$

Here $c(\lambda)$ is the expansion of the determinant and is a polynomial of degree n in λ, called the *characteristic polynomial* of A. Thus

$$c(\lambda) = \lambda^n + c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_1\lambda + c_0$$

The values of λ are precisely the values that satisfy the characteristic equation, and are called the eigenvalues of A.

Example. Find the characteristic equation for the matrix $A = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$.

The method of Faddeev

If the characteristic polynomial of an $n \times n$ matrix A is written as

$$\lambda^n - p_1\lambda^{n-1} - \cdots - p_{n-1}\lambda - p_n$$

then the coefficients $p_1, p_2, \cdots, p_n$ can be computed using

$$p_r = \operatorname{trace} A, \quad (r = 1, 2, \cdots, n)$$

where

$$A_r = \begin{cases} A & (r = 1) \\ AB_{r-1} & (r = 2, 3, \cdots, n) \end{cases}$$

and

$$B_r = A_r - p_r I, \quad \text{where } I \text{ is the } n \times n \text{ identity matrix}$$

The calculations may be checked using the result that

$$B_n = A_n - p_n I \quad \text{must be the zero matrix}$$

1.4.2 Eigenvalues and Eigenvectors

Example. Determine the eigenvalues and eigenvectors for the matrix $A = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$.

1.4.3 Exercises

1. To be written!

1.4.4 Repeated Eigenvalues

Example. Determine the eigenvalues and corresponding eigenvectors of the matrix $A = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}$.

Solution. $\lambda_1 = 4, \lambda_2 = \lambda_3 = 2$

1.5 Numerical Methods

1.6 Reduction to Canonical Form

1.7 Functions of a Matric

1.8 Single Value Decomposition

1.9 State-space Representation

1.10 Solution of the State align

1.11 Engeneering Application: Lyapunov Stability Analysis

1.12 Engineering Application: Capacitor Microphone