Elementary Linear Algebra

Systems of Linear Equations and Matrices

Rules of Matrix Arithmetic

1) $A \pm B = B \pm A$
2) $(A \pm B) \pm C = A \pm (B \pm C)$
3) $A(BC) = (AB)C$
4) $A(B \pm C) = (AB \pm AC)$
5) $(B \pm C)A = BA \pm CA$
6) $a(A \pm B) = aA \pm aB$
7) $ab(C) = a(bC)$
8) $a(BC) = (aB)C = B(aC)$

   A product of invertible matrices is always invertible, and the inverse of the product is the product of the inverse in reverse order.

Elementary Matrices and Method for Finding A^{-1}

Definition > An $n \times n$ matrix is called an *elementary matrix* if it can be obtained from the $n \times n$ identity matrix by performing a single elementary row operation.

Theorem > If the elementary matrix E results from performing a certain row operation on I_m and if A is an $m \times n$ matrix, the product EA is the matrix that results when this same row operation is performed on A.

Example

Consider the matrix

$$A = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix}$$

and consider the elementary matrix

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

which results in adding 3 times the first row of I_3 to the third row. This product EA is then

$$EA = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 9 \end{bmatrix}$$

Remark: Theorem 8 is primarily of theoretical interest and will be used for developing some results about matrices of linear equations.

Theorem > Every elementary matrix is invertible, and the inverse is also an elementary matrix.

Matrices that can be obtained from one another by a finite sequence of elementary row operations are said to be *row equivalent*.

The inverse of a matrix can be found by multiplying it by a sequence of elementary matrices until the identity matrix is created.

$$A^{-1} = E_1 \cdots E_n I_n$$

Further Results on Systems of Equations and Invertibility

If we have a square, invertible matrix and an equation

$$AX = B$$

then we can say

$$X = A^-1B$$

Determinants

The Determinant Function

Elementary Product

Any product of n entries from a matrix A, no two of which come from the same column.

Determinant Function

Let a be a square matrix. The determinant function is denoted by det and we define $det(A)$ to be the sum of all signed elementary products from A.

Evaluating Determinants by Row Reduction

Theorem 1

If A is any square matrix that contains a row of zeros, then $det(A) = 0$

Theorem 2

If A is an $n \times n$ triangular matrix, then $det(A)$ is the product of the entries on the main diagonal; that is; $det(A) = a_{11}a_{22} \cdots a_{nn}$

Properties of the Determinant Function

Theorem 4

If A is any square matrix, then $det(A) = det(A^T)$

In general $det(A + B) \neq det(A) + det(B)$, however;

Theorem 5

If A and B are square matrices of the same size, then $det(AB) = det(A)det(B)$

Theorem 6

A matrix is invertible iff $det(A) \neq 0$

Cofactor Expansion; Cramer's Rule

Definition: minor entry and cofactor

If A is a square matrix, then the minor entry a_{ij} is denoted by M_{ij} and is defined to be the determent few the submatrix that remains after the i_{th} row and j_{th} column are deleted from A. The number $(-1)^{i+j}M_{ij}$ is denoted by C_{ij} and is called the cofactor of entry a_{ij}

$$det(A) = a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31}$$

Definition

Any $n \times n$ matrix made by the cofactors of a_{ij} is called the matrix of cofactors from A. The transpose of this matrix is called the **adjoint** of A and is denoted $adj(A)$.

Theorem 8

If A in an invertible matrix then

$$A^{-1} = \frac{1}{det(A)} adj(A)$$

Determinants can be computed by multiplying the entries in the first column of A by their cofactors and adding the resulting products. This method is called **cofactor expansion**.

Vector Spaces

Basis and Dissension

You can show a system of linear equations are linearly independent and span R^n by showing the matrix has a non-zero determinant.

Row and Column of A Matrix; Rank

Theorem 12

If A is any matrix, then the row space and column space have the same dimension (**rank**)

Theorem 13

All of the following statements are equivalent

1. A is invertible
2. $Ax = 0$ has only the trivial solution
3. A is row equivalent to I
4. $Ax = b$ is consistent for every $n \times 1$ matrix b
5. $det(A) \neq 0$
6. A has rank n
7. The row vectors of A are linear independent
8. The column vectors of A are linearly independent

Theorem 14

A system of linear equations $Ax = b$ is consistent iff b is in the column space of A.

Inner Product Spaces

Theorem 15

If u and v are vectors in an inner product space V, then

$$< u, v > \leq < u, u > < v, v >$$

Length and Angle In Inner Product Spaces

Norm: $||u|| =< u, u >^{1/2}$

Distance: $d(u, v) = ||u - v||$

Angle: $cos(\theta) = \frac{<u,v>}{||u|| \, ||v||} \ \ and \ \ 0 \leq \theta \leq \pi$

Theorem 17 (Generalized Theorem of Pythagoras): If u and v are orthogonal vectors in an inner product space, then

$$||u + v||^2 = ||u||^2 + ||v||^2$$

Orthonormal Bases; Gram-Schmidt Process

Definition

A set of vectors in an inner product is called an *orthogonal set* if any two distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm 1 is called *orthonormal*.

Gram-Schmidt Process

In order to find a orthonormal basis for $\mathbb{R}^n$ consider the following

1. $v_1 = \frac{u_1}{||u_1||}$
2. $v_2 = \frac{u_2 - <u_2,v_1>v_1}{||u_2 - <u_2,v_1>v_1||}$
3. $v_3 = \frac{u_3 - <u_3,v_1>v_1 - <u_3,v_2>v_2}{||u_3 - <u_3,v_1>v_1 - <u_3,v_2>v_2||}$
4. $v_4 = \frac{u_4 - <u_4,v_1>v_1 - <u_4,v_2>v_2 - <u_4,v_3>v_3}{||u_4 - <u_4,v_1>v_1 - <u_4,v_2>v_2|| - <u_4,v_3>v_3}$
5. $\vdots$

Theorem 23 (Best Approximation Theorem):

If W is a finite dimensional subspace of an inner product space V, and if u is a vector in V, then $proj_w u$ is the best approximation to u from w in the sense that

$$||u - proj_W u|| < ||u - w||$$

for every vector w in W different from $proj_W u$.

Coordinates; Change of Basis

Theorem 24: If $S = v_1, v_2, \cdots, v_n$ is a basis for a vector space V, then every vector in v in V can be expressed in the form $v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$ in exactly one way.

The **coordinate vector** of v relative to S is denoted by $(v)_S$ and is the vector in $\mathbb{R}^n$ defined by:

$$(v)_S = (c_1, c_2, \cdots, c_n)$$

The **coordinate matrix** of v relative to S is denoted by $[v]_S$ and is the $n \times 1$ matrix defined by:

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

If we change the basis for a vector space V from some old basis B to some new basis B' then the old coordinate matrix $[v]_B$ of a vector v is related to the new coordinate matrix $[v]_{B'}$ by the equation

$$[v]_B = P[v]_{B'}$$

where the columns of P are the coordinate matrices of the old basis vectors relative to the new basis. P is formally known as the **transition matrix** from B to B'.

Theorem 26:

If P is the transition matrix from basis B to basis B' then:

1. P is invertible
2. P^{-1} is the transition matrix from B' to B

Theorem 27:

If P is the transition matrix from one orthonormal basis to another for an inner product space then

$$P^{-1} = P^T$$

Definition:

A square matrix with the property

$$A^{-1} = A^T$$

is said to be an orthogonal matrix.

Theorem 28:

1. A is orthogonal
2. The row vectors of A form an orthonormal set in R^n with the Euclidean inner product
3. The column vectors of A for an orthonormal set in R^n with the Euclidean inner product

Linear Transformations

Introduction to Linear Transformations

If V and W are vector spaces and F is a function that associates a unique vector in W with each vector in V, we say F **maps** V into W and write $F : V \to W$. Furthermore, if F associates the vector w with the vector v, we write $w = F(v)$ and say that w is the **image** of v under F.

Definition:

> If $F : V \to W$ is a function from the vector space V into the vector space W, then F is called a **linear transformation** if:

1. $F(v + v) = F(u) + F(v)$
2. $F(ku) = kF(u)$

Properties of Linear Transformations; Kernel and Range

Definition

> if $T : V \to W$ is a linear transformation, then the set of vectors in V that T maps into 0 is called the **kernel** (or **nullspace**) of T; it is denoted by $ker(T)$. The set of all vectors in W that are images under T of at least one vector in V is called the **range** of T; it is denoted by $R(T)$.

Theorem 2

> If $T : V \to W$ is a linear transformation then:

1. The kernel of T is a subspace of V
2. The range of T is a subspace of W

Definition

> if $T : V \to W$ is a linear transformation, then the dimension of the range of T is called the **rank of** T and the dimension of the kernel is called the **nullity of** T.

Theorem 3 (Dimensions Theorem):

> If $T : V \to W$ is a linear transformation from an n-dimensional vector space V to a vector space W, then

$$rank(T) + ker(T) = n$$

> In the special case where $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ is multiplication by an $m \times n$ matrix A, the dimensional theorem yields the following:

$$nullity\ of\ T = n - rank(T) = number\ of\ columns(T) - rank(T)$$

Theorem 4

> If A is an $m \times n$ matrix then the dimension of the solution space of $Ax = 0$ is

$$n - rank(A)$$

Matrices of Linear Transformations

If V and W are any finite dimension vector space, then with some ingenuity, any linear transformation $T : V \to W$ can be regarded as a matrix transformation. The basic idea is to choose bases for V and W and to work with the coordinate matrices relative to these bases rather than with the vectors themselves. To be specific, suppose V is n-dimensional and W is m-dimensional. If we choose bases B and B' then for each x in V, the coordinate (column) matrix $[x]_b$ will be a vector in $\mathbb{R}^n$ and the coordinate matrix $[T(x)]_{B'}$ will be some vector in $\mathbb{R}^m$. **Thus in the process of mapping x into $T(x)$, the linear transformation T "generates" a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ by sending $[x]_B$ into $[T(x)]_{B'}$.**

$$A[x]_B = [T(x)]_{B'}$$

To find this transform you can follow three steps:

1. Compute the coordinate matrix $[x]_B$
2. Multiply $[x]_B$ on the left by A to produce $[T(x)]_{B'}$.
3. Reconstruct $T(x)$ from its coordinate matrix $[T(x)]_{B'}$

Similarity

Definition

> If A and B are square matrices, we say that B is similar to A if there is an invertible matrix P such that $B = P^{-1}AP$

Eigenvalues, Eigenvectors

Eigenvalues and Eigenvectors

Eigenvalues

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

Eigenvectors

$$Ax = \lambda x$$

Where x is the eigenvector.

Diagonalization

Definition

> A square matrix is **diagonalizable** if there is an invertible matrix P such that $P^{-1}AP$ is diagonal; the matrix P is said to diagonalize A.

The following procedure will diagonalize a matrix:

1. Find n linearly independent eigenvectors of A, $p_1, p_2, \cdots$
2. For the matrix P having $p_1, p_2, \cdots$ as its column vectors
3. The matrix $P^{-1}AP$ will then be diagonal with $\lambda_1, \lambda_2, \cdots, \lambda_n$ as its successive diagonal entries, where λ_i is the eigenvalue corresponding to $p_i,\ i = 1, 2, 3, \cdots, n$

Orthogonal Digitalization; Symmetric Matrices

Definition

> A square matrix A is called **orthogonally diagonalizable** if there is an orthogonal matrix P such that $P^{-1}AP$ is diagonal; that matrix P is said to **orthogonally diagonalize** A.

An orthogonally diagonalizable matrix is **symmetric**. From this we can say:

Theorem

> If A is symmetric, then eigenvectors from different eigenspaces are orthogonal.

As a consequence, the following procedure will find an orthonormal diagonalized matrix

1. Find a basis for each eigenspace of A
2. Apply the Gram-Schmidt process to each of these basis to obtain and orthonormal basis from each eigenspace
3. From the matrix P whose columns are the basis vectors constructed in step 2; this matrix orthogonally diagonalizes A

References

- Elementary Linear Algebra - Howard Anton