ERROR-CORRECTING CODES

W. Wesley Peterson

The M.I.T. Press
Massachusetts Institute of Technology
Cambridge, Massachusetts

In any space, the number of linearly independent vectors that span the space is called the dimension of the space. A set of n linearly independent vectors spanning an n-dimensional vector space is called a basis of the space. It follows from Theorem 2.7 that every set of more than n vectors in an n-dimensional vector space is linearly dependent. It follows from Theorem 2.6 that no set of fewer than n vectors can span an n-dimensional space.

Theorem 2.8. If V is an n-dimensional vector space, any set of n linearly independent vectors in V is a basis for V.

Proof: Let $v_1, v_2, \ldots, v_n$ be a set of linearly independent vectors in V. If they do not span V, there must be some vector v in V that is not a linear combination of $v_1, v_2, \ldots, v_n$. Then the set $v, v_1, v_2, \ldots, v_n$ of $n+1$ vectors in V is linearly independent. This contradicts Theorem 2.6, and therefore $v_1, v_2, \ldots, v_n$ must span V.

QED

Theorem 2.9. If a vector space V_1 is contained in a vector space V_2 and they have the same dimension n, they are equal.

Proof: A basis for V_1 is a set of n linearly independent vectors in V_2. Therefore every vector in V_2 is also in V_1.

QED

An inner product or dot product of two n-tuples is a scalar and is defined as follows:

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = a_1 b_1 + \ldots + a_n b_n$$

It is easily verified that $u \cdot v = v \cdot u$ and that $w \cdot (u + v) = w \cdot u + w \cdot v$. If the inner product of two vectors is zero, they are said to be orthogonal.

2.6. Matrices

The purpose of this section is to outline the parts of matrix theory that apply to the codes studied in the next three chapters. For the most part, proofs are given, but this can hardly serve as more than a review of the necessary parts of matrix theory.

An $n \times m$ matrix is an ordered set of nm elements in a rectangular array of n rows and m columns:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} = [a_{ij}]$$

The elements of a matrix may in general be elements of any ring, but in this book only matrices with elements in a field find application. The n rows may be thought of as n m-tuples or vectors, and similarly, the m columns may be thought of as vectors. The set of elements a_{ii} for which the column number and row number are equal is called the main diagonal.

The row space of an $n \times m$ matrix M is the set of all linear combinations of row vectors of M. They form a subspace of the vector space of m-tuples. The dimension of the row space is called the row rank. Similarly, the set of all linear combinations of column vectors of the matrix forms the column space, whose dimension is called the column rank.

There is a set of elementary row operations defined for matrices:

1. Interchange of any two rows.
2. Multiplication of any row by a nonzero field element.
3. Addition of any multiple of one row to another.

The inverse of each elementary row operation is clearly an elementary row operation of the same kind.

Theorem 2.10. If one matrix is obtained from another by a succession of elementary operations, both matrices have the same row space.

Proof: If the theorem is true for each elementary row operation, it will clearly be true for a succession. It is obviously true of row operations 1 and 2. Suppose that the matrix M' is obtained from the matrix M by type 3 elementary row operation. Then, since the altered row of M' is a linear combination of two rows of M, any linear combination of rows of M' is also a linear combination of rows of M, so the row space of M' is contained in the row space of M. But M can be obtained from M' by the inverse operation, which is again an operation of type 3, so the row space of M must be contained in the row space of M''. Therefore they are equal.

QED

Elementary row operations can be used to simplify a matrix and put it in a standard form. The form, called echelon canonical form, is as follows:

1. Every leading term of a nonzero row is 1.
2. Every column containing such a leading term has all its other entries zero.
3. The leading term of any row is to the right of the leading term in every preceding row. All zero rows are below all nonzero rows.

The procedure is essentially the same as that used in solving linear equations by elimination of one variable at a time. It is

best illustrated by an example. Consider the following matrix with real numbers as elements:

$$\begin{bmatrix} 0 & 0 & 2 & 2 & 0 & 2 \\ 2 & 2 & 6 & 8 & 4 & 8 \\ 1 & 1 & 5 & 6 & 2 & 5 \\ 1 & 1 & 3 & 4 & 2 & 7 \end{bmatrix}$$

To simplify the matrix, the first step would be to locate the first column with a nonzero element, interchange rows if necessary to place a nonzero element in the first row, and multiply the row by the inverse of that element to give a leading 1 . Interchanging rows 1 and 2 and dividing by 2 give

$$\begin{bmatrix} 1 & 1 & 3 & 4 & 2 & 4 \\ 0 & 0 & 2 & 2 & 0 & 2 \\ 1 & 1 & 5 & 6 & 2 & 5 \\ 1 & 1 & 3 & 4 & 2 & 7 \end{bmatrix}$$

The next step is to subtract a multiple of the first row from each other row to make the rest of the column corresponding to the leading element in the first row 0 :

$$\begin{bmatrix} 1 & 1 & 3 & 4 & 2 & 4 \\ 0 & 0 & 2 & 2 & 0 & 2 \\ 0 & 0 & 2 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

Then, with disregard of the first row, again the first column with a nonzero element is located, and rows are interchanged if necessary to place a nonzero element in this column in the second row. The row is next multiplied by the inverse of its leading element to give a leading 1 . This is accomplished in the above matrix by dividing the second row by 2 . Then the appropriate multiple of this row is subtracted from each other row to make all the other entries 0 in the column of the leading element of the second row. This yields

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

One more step in the process yields

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This process will always result in a matrix in echelon canonical form.

The nonzero rows of a matrix in echelon canonical form are linearly independent, and thus the number of nonzero rows is the dimension of the row space. It can be shown that there is only one matrix in echelon canonical form for any given row space.

If all the rows of an $n \times n$ matrix are linearly independent, the matrix is said to be nonsingular. When such a matrix is put in echelon canonical form, there must still be n linearly independent rows, and thus every row must contain a 1. This can occur only if it has 1's on the main diagonal and 0's elsewhere. Such a matrix is called an identity matrix and denoted I. Thus any nonsingular matrix can be transformed into an identity matrix by elementary row operations.

The transpose of an $n \times m$ matrix M is an $m \times n$ matrix, denoted M^T, whose rows are the columns of M, and thus whose columns are the rows of M. The transpose of $[a_{ij}]$ is $[a_{ji}]$.

Two $n \times m$ matrices can be added, element by element:

$$[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$$

With this definition it is easily verified that matrices form an Abelian group under addition.

An $n \times k$ matrix $[a_{ij}]$ and a $k \times m$ matrix $[b_{ij}]$ can be multiplied to give an $n \times m$ product matrix $[c_{ij}]$ by the rule

$$c_{ij} = \sum_{\ell=1}^{k} a_{i\ell} b_{\ell j}$$

It can be verified by direct calculation that with this definition matrix multiplication satisfies the associative law, and multiplication and addition satisfy the distributive law.

The element c_{ij} of the product is the inner product of the i^{th} row of $[a_{ij}]$ by the j^{th} column of $[b_{ij}]$. Also the i^{th} row vector of the product $[c_{ij}]$ is a linear combination of the row vectors of $[b_{\ell j}]$ with the coefficient $a_{i\ell}$ on the ℓ^{th} row. Similarly the columns of the product are linear combinations of the column vectors of $[a_{i\ell}]$.

Multiplying an $n \times m$ matrix M on the left by a matrix P that has one 1 in each row and each column and all the rest of the elements 0 simply permutes the rows of the matrix M, and any