2. MATRIX AND INDICIAL NOTATION

2.1 Initial Comments

It is essential to know basic matrix and indicial notation. Here we provide a brief review of matrix notation. Since we will be working with only three-dimensional vectors and matrices, it is not necessary at this time to provide a comprehensive survey of matrix theory. However, these relations must be at your fingertips.

In the second part, indicial notation is introduced. In many cases, indicial notation has a direct correlation with matrix notation, and when these correlations exist we switch from one notation to the other whenever one is more convenient than the other. However, there are situations where indicial notation has no direct analogue with matrix notation.

2.2 Matrix Notation

<u>Vectors</u>

A set of three numbers can be ordered within a mathematical vector in row and column form as follows:

$$\langle a\rangle = \left\langle \begin{matrix} a_1 & a_2 & a_3 \end{matrix} \right\rangle \qquad \{a\} = \left\{ \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix} \right\} \tag{2-1}$$

The transpose operator converts one form to the other

$$\langle a\rangle^T = \{a\} \qquad \{a\}^T = \langle a\rangle \tag{2-2}$$

The inner product of two vectors is defined to be

$$\langle a\rangle\{b\} = a_1 b_1 + a_2 b_2 + a_3 b_3 \tag{2-3}$$

The inner product of a vector with itself is the square of the magnitude

$$\left|\{a\}\right|^2 = \langle a\rangle\{a\} \tag{2-4}$$

A unit vector formed from $\{a\}$ is

$$\{u\}^a = \frac{1}{\left|\{a\}\right|}\{a\} \tag{2-5}$$

If $\{a\}$ is multiplied by a scalar, α , the result is

$$\alpha\{a\} = \left\{ \begin{matrix} \alpha a_1 \\ \alpha a_2 \\ \alpha a_3 \end{matrix} \right\} \tag{2-6}$$

<u>Square 3x3 Matrices</u>

A set of nine numbers is ordered within a matrix as follows:

$$[a] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{2-7}$$

in which the first index of each component denotes the row and the second index identifies the column. Two alternative ways to denote the components of the matrix is to use row and column vectors as follows:

$$[a] = \begin{bmatrix} \{a\}_1 & \{a\}_2 & \{a\}_3 \end{bmatrix} = \begin{bmatrix} \langle a \rangle_1 \\ \langle a \rangle_2 \\ \langle a \rangle_3 \end{bmatrix} \tag{2-8}$$

where

$$\{a\}_1 = \begin{Bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{Bmatrix} \quad \text{etc.} \qquad \langle a \rangle_1 = \left\langle \; a_{11} \quad a_{12} \quad a_{13} \; \right\rangle \quad \text{etc.} \tag{2-9}$$

The transpose of a matrix is the matrix obtained by interchanging rows and columns, or

$$[a]^T = \begin{bmatrix} \left\langle \{a\}_1^T \right\rangle \\ \left\langle \{a\}_2^T \right\rangle \\ \left\langle \{a\}_3^2 \right\rangle \end{bmatrix} = \begin{bmatrix} \left\{ \langle a \rangle_1^T \right\} & \left\{ \langle a \rangle_2^T \right\} & \left\{ \langle a \rangle_3^T \right\} \end{bmatrix} \tag{2-10}$$

The multiplication of a matrix by a column vector is a column vector with components as follows:

$$[a]\{v\} = \{b\} \qquad \{b\} = \begin{bmatrix} \langle a \rangle_1 \\ \langle a \rangle_2 \\ \langle a \rangle_3 \end{bmatrix} \begin{Bmatrix} v_1 \\ v_2 \\ v_3 \end{Bmatrix} = \begin{Bmatrix} \langle a \rangle_1 \{v\} \\ \langle a \rangle_2 \{v\} \\ \langle a \rangle_3 \{v\} \end{Bmatrix} \tag{2-11}$$

or, equivalently, the resulting vector can be considered as a sum of column vectors:

$$[a]\{v\} = \{b\} \qquad \{b\} = \begin{bmatrix} \{a\}_1 & \{a\}_2 & \{a\}_3 \end{bmatrix} \begin{Bmatrix} v_1 \\ v_2 \\ v_3 \end{Bmatrix} = v_1\{a\}_1 + v_2\{a\}_2 + v_3\{a\}_3 \tag{2-12}$$

The result of pre-multiplying a matrix by a row vector is a row vector with components as follows:

$$\langle u\rangle[a]=\langle c\rangle \qquad \{c\}=\left\langle\; u_1 \quad u_2 \quad u_3 \;\right\rangle\left[\; \{a\}_1 \quad \{a\}_2 \quad \{a\}_3 \;\right]$$
$$=\left\langle\; \langle u\rangle\{a\}_1 \quad \langle u\rangle\{a\}_2 \quad \langle u\rangle\{a\}_3 \;\right\rangle \tag{2-13}$$

or, equivalently, the resulting vector can be considered as a sum of row vectors:

$$\langle u\rangle[a]=\langle c\rangle \qquad \{c\}=\left\langle\; u_1 \quad u_2 \quad u_3 \;\right\rangle\begin{bmatrix} \langle a\rangle_1 \\ \langle a\rangle_2 \\ \langle a\rangle_3 \end{bmatrix}$$
$$=u_1\langle a\rangle_1+u_2\langle a\rangle_2+u_3\langle a\rangle_3 \tag{2-14}$$

The product of two matrices is a matrix with components as follows:

$$[a][b]=[c] \qquad [c]=\begin{bmatrix} \langle a\rangle_1\{b\}_1 & \langle a\rangle_1\{b\}_2 & \langle a\rangle_1\{b\}_3 \\ \langle a\rangle_2\{b\}_1 & \langle a\rangle_2\{b\}_2 & \langle a\rangle_2\{b\}_3 \\ \langle a\rangle_3\{b\}_1 & \langle a\rangle_3\{b\}_2 & \langle a\rangle_3\{b\}_2 \end{bmatrix} \tag{2-15}$$

The transpose of a product of two matrices, or of a matrix with a vector, is the product of the transposes in reverse order. Examples are:

$$\begin{array}{llll} [a]\{u\}=\{v\} & \text{Take transpose } \{[a]\{u\}\}^T=\{v\}^T & \Rightarrow & \langle u\rangle[a]^T=\langle v\rangle \\ \langle w\rangle[a]=\langle z\rangle & \text{Take transpose } \langle\langle w\rangle[a]\rangle^T=\langle z\rangle^T & \Rightarrow & [a]^T[w]=\{z\} \\ [a][b]=[c] & \text{Take transpose } [[a][b]]^T=[c]^T & \Rightarrow & [b]^T[a]^T=[c]^T \end{array} \tag{2-16}$$

The trace of a matrix, denoted $tr[a]$, is defined to be the sum of the diagonal terms, or

$$tr[a]=a_{11}+a_{22}+a_{33} \tag{2-17}$$

Matrices of Arbitrary Size

In general a matrix may not be square and have arbitrary numbers of rows and columns. The notation $[M]_{mxn}$ denotes a matrix with m rows and n columns. The product of two matrixes $[M]_{mxn}$ and $[N]_{rxs}$ is defined only when $r=n$. The form for each component in the product is as follows:

$$[P]_{mxs} = [M]_{mxn}[N]_{nxs} \qquad P_{ij} = \sum_{k=1}^{n} M_{ik}N_{ks} \quad i = 1,\cdots,m \quad \mathrm{j} = 1,\cdots,s \tag{2-18}$$

In this context a column vector with 3 components can be considered a matrix with 3 rows and one column. Similarly a row vector with 3 components is a matrix with one row and 3 columns. With the use of (2-18), the product of a row vector and a column vector is the matrix

$$\{u\}\langle v\rangle = \begin{bmatrix} u_1v_1 & u_1v_2 & u_1v_3 \\ u_2v_1 & u_2v_2 & u_2v_3 \\ u_3v_1 & u_3v_2 & u_3v_3 \end{bmatrix} \tag{2-19}$$

Recall that $\langle u\rangle\{v\}$ is a scalar so it is essential that care be taken in specifying the sequence of multiplication for the product of two vectors.

Particular Matrices

A matrix is symmetric if $[a]=[a]^T$ and skew-symmetric if $[a]=-[a]^T$. Any matrix can be decomposed into a sum of symmetric and skew-symmetric parts:

$$[a]=[a]_{sy}+[a]_{sk} \quad \text{where} \quad [a]_{sy}=\frac{1}{2}\Big[[a]+[a]^T\Big] \qquad [a]_{sk}=\frac{1}{2}\Big[[a]-[a]^T\Big] \tag{2-20}$$

Show that $[a]_{sy}$ contains six independent components, and the diagonal components of $[a]_{sk}$ are zero and the off-diagonal terms are defined by three independent components.

The identity matrix is defined to be

$$[I] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2-21}$$

The product of the identity with a matrix is the matrix itself.

Inverse of a Matrix and Determinants

For a given matrix $[a]$ the inverse $[a]^{-1}$ satisfies

$$[a][a]^{-1}=[I] \qquad [a]^{-1}[a]=[I] \tag{2-22}$$

If the matrix $[a]$ is orthogonal, then

$$[a]^{-1}=[a]^T \tag{2-23}$$

This is an extraordinarily convenient relation that we will employ extensively in conjunction with transformation matrices. If a matrix is not orthogonal, then the procedure for obtaining an inverse is not elementary. One approach, which is satisfactory for matrices up to size 3x3, is to use determinants. For one, two and three-dimensional square matrices the determinants are

$$\left|a\right| = a \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{12} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{31} \\ -a_{12}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{22}a_{33} \end{matrix} \tag{2-24}$$

To obtain an inverse, we first define a submatrix $\left[A_{kl}\right]$ as the matrix obtained by deleting the k'th row and l'th column of a matrix $\left[a\right]$. For example

$$\left[A_{23}\right] = \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} \tag{2-25}$$

The components of a co-factor matrix $\left[A^{cf}\right]$ are defined to be

$$A_{ij}^{cf} = (-1)^{i+j}\left\|\left[A_{ij}\right]\right\| \tag{2-26}$$

Note that to obtain the cofactor matrix, the determinants of nine submatrices must be obtained. Then the inverse matrix involves the transpose of the cofactor matrix (called the adjoint matrix) as follows:

$$\left[a\right]^{-1} = \frac{1}{\left\|\left[a\right]\right\|}\left[A^{cf}\right]^T \tag{2-27}$$

This equation requires a great deal more effort than that of merely taking the transpose of an orthogonal matrix.

Similarly to the property for transposes, the inverse of a product of matrices is the product of the inverses in reverse order:

$$\left[\left[a\right]\left[b\right]\right]^{-1} = \left[b\right]^{-1}\left[a\right]^{-1} \tag{2-28}$$

The order of taking the inverse and a transpose is immaterial:

$$\left[\left[a\right]^{-1}\right]^T = \left[\left[a\right]^T\right]^{-1} \tag{2-29}$$

The determinant of a product of matrices is the product of the determinants:

$$\left\|\left[a\right]\left[b\right]\right\| = \left\|\left[a\right]\right\|\left\|\left[b\right]\right\| \tag{2-30}$$

Derivative of the Determinant

Later we will need the derivative of the determinant with respect to a component of the matrix. See the extended notes that give the derivation for the following equation that also involves the components of the cofactor matrix:

$$\frac{\partial \|[a]\|}{\partial a_{ij}} = A_{ij}^{cf} \qquad i,j = 1,2,3 \tag{2-31}$$

The Linear Algebraic Problem

The linear algebraic problem is one of finding $\{x\}$ from a set of linear algebraic equations when the coefficient matrix $[a]$ and the force vector $\{b\}$ are known:

$$[a]\{x\} = \{b\} \tag{2-32}$$

One method is to obtain the inverse $[a]^{-1}$ and the solution is

$$\{x\} = [a]^{-1}\{b\} \tag{2-33}$$

A second method is to use Cramer's rule

$$\{x\} = \frac{1}{D}\{X\} \qquad D = \|[a]\| \qquad X_1 = \left\| \begin{bmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{bmatrix} \right\| \quad \text{etc.} \tag{2-34}$$

Other methods that are used for linearly solving larger problems involve LU and QR decompositions ou f $[a]$.

The Eigenproblem

If a vector $\{x\}$ and a scalar λ exist such that

$$[a]\{x\} = \lambda\{x\} \tag{2-35}$$

then $\{x\}$ and λ are called an eigenvector and an eigenvalue, respectively, of $[a]$. Together $(\lambda, \{x\})$ are called an eigenpair. The eigenproblem is to find the eigenpairs of a given $[a]$. This problem will be considered in depth in the chapter on second-order tensors.

Other Matrix Properties

The above represents the bare bones that you have to know for this course. Other properties for matrices and determinants and some derivations are provided in the extended notes. Of course more complete coverage is available in books on matrix theory.

2.3 Indicial Notation

Single Index

If a letter, u, is associated with a single index (called a "free" index) such as u_i, or u_j, or u_k, where the index assumes the values 1, 2 and 3, then any one of these representations implies (using the symbol $\Rightarrow$) the same set of three numbers:

$$u_i \Rightarrow (u_1, u_2, u_3) \qquad u_j \Rightarrow (u_1, u_2, u_3) \qquad \text{the same set} \tag{2-36}$$

It is the letter "u" that identifies the set; the index indicates that it is a set of three numbers. A value of an index identifies the particular member of the set. We cannot use an equal sign because one rule of indicial notation is that an equation must have the same free index for each collection of terms within an equation.

We may want to convert to matrix notation. The set u_i can be interpreted as a column vector, or as a row vector, whichever is most convenient for the problem at hand. We represent these alternative representations as follows:

$$u_i \Rightarrow \{u\} \qquad u_i \Rightarrow \langle u \rangle \tag{2-37}$$

Two Indices

If a letter "a" is associated with two free indices, then any one of the notations a_{ij} or a_{AB} or $a_{\alpha\beta}$ denotes a set of nine members that we often represent as a matrix:

$$a_{ij} \Rightarrow [a] \qquad a_{AB} \Rightarrow [a] \qquad a_{\alpha\beta} \Rightarrow [a] \qquad \text{a single set of numbers} \tag{2-38}$$

with the first index denoting the row and the second the column. Again, we cannot use an equal sign because the right side does not contain the free indices i and j. The transpose of a two-index set is obtained by interchanging the indices:

$$\left(a_{ij}\right)^T = a_{ij}^T = a_{ji} \tag{2-39}$$

Now if we take the transpose of the transpose, we obtain the original set:

$$a_{ji}^T = a_{ij} \tag{2-40}$$

Multiple Indices

In a similar manner, we let f_{ijk} and g_{ABCD} represent sets of 27 and 81 members, respectively. There is no straightforward way of representing these sets with matrix notation although we can use constructions such as

$$a_{jk} = f_{1jk} \qquad b_{jk} = f_{2jk} \qquad \mathrm{c}_{jk} = f_{3jk} \tag{2-41}$$

for convenience. In general, if the number of free indices is m, then the set contains 3^m members.

Summation Convention

If two indices are repeated, then a summation of terms is implied. Examples are:

$$u_i v_i = u_k v_k = u_\alpha v_\alpha = \sum_{A=1}^{3} u_A v_A = u_1 v_1 + u_2 v_2 + u_3 v_3 \tag{2-42}$$

A pair of repeated indices may replace any other pair of repeated indices. Since there is no free index, the result is a scalar. Therefore, we can actually use an equal sign to equate the summation convention with vector multiplication

$$u_k v_k = \langle u \rangle \{v\} = \langle v \rangle \{u\} \tag{2-43}$$

For a set with two indices, the summation convention can be used to obtain the scalar

$$c = a_{ii} = a_{AA} = a_{\alpha\alpha} = a_{11} + a_{22} + a_{33} = tr[a] \tag{2-44}$$

No index can be repeated more than twice unless the summation symbol is used, e.g.,

$$c = \sum_{j=1}^{3} a_{jj} v_j \tag{2-45}$$

Indicial Equations

Terms with indices can always be multiplied together. The order of the identifier is irrelevant but the order of the indices is important. Each product of terms must contain the same free indices. Therefore the equation $u_i = v_j$ is not allowed. Instead $u_i = v_i$ is valid and represents a set of three equations

$$u_i = v_i \quad \Rightarrow \quad u_1 = v_1 \quad u_2 = v_2 \quad u_3 = v_3 \tag{2-46}$$

Any index can be used as long as the same one occurs on both sides of an equation, i.e., (2-46) is identical to

$$u_j = v_j \quad u_A = v_A \quad u_\alpha = v_\alpha \tag{2-47}$$

Another example is

$$a_{ij} = b_{ij} \text{ or } a_{AB} = b_{AB} \quad \Rightarrow \text{ a set of nine eqs. } a_{11} = b_{11}, \text{a}_{12} = b_{12} \text{ etc.} \tag{2-48}$$

Other examples of indicial equations are:

$$\begin{array}{ll} u_i v_j w_k = X_{ijk} & \text{a set of 27 equations} \\ \text{u}_i F_{mmj} = a_{ij} & \text{a set of 9 equations that include a summation} \\ \text{Y}_{ijkl} X_{kl} = R_{ij} & \text{a set of 9 equations that include a double summation} \end{array} \tag{2-49}$$

On the left side of these equations the identifiers can appear in any order, i.e., $u_i F_{mmj} = F_{mmj} u_i$.

Now we look at special cases that have corresponding matrix representations. First consider

$$u_i = v_j a_{ij} = a_{ij} v_j \tag{2-50}$$

The identifiers in the product can be written in any order. The second form has the summation indices adjacent to each other and this suggests that we have the matrix correspondence of

$$u_i = b_{ij} v_j \qquad \Rightarrow \qquad \{u\} = [b]\{v\} \tag{2-51}$$

Now suppose we consider

$$w_A = z_B a_{BA} = a_{BA} z_B \tag{2-52}$$

The first form has adjacent summation indices so we conclude that

$$w_A = z_B a_{BA} \qquad \Rightarrow \qquad \langle w \rangle = \langle z \rangle [a] \tag{2-53}$$

Suppose we rewrite (2-52) using a transpose. The result is

$$w_A = a^T_{AB} z_B \qquad \Rightarrow \qquad \{w\} = [a]^T \{z\} \tag{2-54}$$

which is identical to the result obtained by taking the transpose of the vector equation in (2-53).

Next we consider the indicial analogue of matrix multiplication as follows:

$$\begin{aligned} &c_{ij} = a_{ik} b_{kj} \qquad \Rightarrow \qquad [c] = [a][b] \\ &c_{ij} = a_{ik} b_{jk} = a_{ik} b^T_{kj} \qquad \Rightarrow \qquad [c] = [a][b]^T \\ &c = a_{ij} b_{ji} \qquad \Rightarrow \qquad \mathrm{c} = tr\big[[a][b]\big] \\ &c = a_{ij} b_{ij} = a_{ij} b^T_{ji} \qquad \Rightarrow \qquad \mathrm{c} = tr\Big[[a][b]^T\Big] \end{aligned} \tag{2-55}$$

Special Symbols

The Kronecker delta, δ_{ij} is defined to be 1 if $i = j$ and 0 if $i \neq j$ with the following implications:

$$\begin{aligned} &\delta_{ij} \qquad \Rightarrow \qquad [I] \\ &\delta_{ij} u_i = u_j \qquad \delta_{ij} b_{ik} = b_{jk} \\ &f_{ijk} \delta_{jk} = f_{ijj} = f_{jkk} \\ &a_{ij} \delta_{ij} = a_{ii} = a_{jj} = tr[a] \\ &d_{ijk} a_{mn} \delta_{kn} = d_{ijk} a_{mk} = d_{ijn} a_{mn} \end{aligned} \tag{2-56}$$

Note that every product of terms in any one equation has the same free indices. If one of indices in the Kronecker delta is summed with a product term, the Kronecker delta vanishes and the repeated index in the product term is replaced with the other index from the Kronecker delta.

The alternating symbol ε_{ijk} is defined as follows:

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if i,j,k form a positive permutation} \\ -1 & \text{if i,j,k form a negative permutation} \\ 0 & \text{if two or more indices are the same} \end{cases} \tag{2-57}$$

or

$$\begin{aligned} &\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1 \\ &\varepsilon_{132} = \varepsilon_{321} = \varepsilon_{213} = -1 \\ &\varepsilon_{113} = 0 \quad \varepsilon_{122} = 0 \quad \text{etc.} \end{aligned} \tag{2-58}$$

Any time two indices are the same, the symbol is zero.

It follows that permuting the indices forward does not change the value and interchanging any two indices changes the sign, or

$$\begin{aligned} &\varepsilon_{ijk} = \varepsilon_{jki} = \varepsilon_{kij} \\ &\varepsilon_{ikj} = -\varepsilon_{ijk} \quad \varepsilon_{kji} = -\varepsilon_{ijk} \quad \varepsilon_{jik} = -\varepsilon_{ijk} \end{aligned} \tag{2-59}$$

Note that many of these equations are simply zero equals zero.

An example of the use of the alternating symbol is the following equation:

$$w_k = \varepsilon_{ijk} u_j v_k \quad \Rightarrow \quad w_1 = u_2 v_3 - u_3 v_2 \quad w_2 = u_3 v_1 - u_1 v_3 \quad w_3 = u_1 v_2 - u_2 v_1 \tag{2-60}$$

an equation that is used extensively to obtain the components of the cross product of two vectors.

Alternating Symbol-Kronecker Delta Identity

The alternating symbol-Kronecker delta identity is

$$\varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} \tag{2-61}$$

It follows that

$$\begin{aligned} &\varepsilon_{ijk}\varepsilon_{ijm} = \delta_{jj}\delta_{km} - \delta_{jm}\delta_{kj} = 3\delta_{km} - \delta_{km} = 2\delta_{km} \\ &\varepsilon_{ijk}\varepsilon_{ijk} = 2\delta_{kk} = 6 \end{aligned} \tag{2-62}$$

Alternating Symbol-Determinant Identity

Another important identity involves the determinant of a matrix $[a]$ with components a_{ij} as follows:

$$\varepsilon_{ijk} a_{il} a_{jm} a_{kn} = \varepsilon_{lmn} \left|[a]\right| \tag{2-63}$$

With the use of (2-62) it is easily shown that

$$\varepsilon_{lmn}\varepsilon_{ijk}a_{il}a_{jm}a_{kn} = 6\left|[a]\right| \tag{2-64}$$

These equations are often used in theoretical derivations and not as a procedure for obtaining a determinant.

2.4 Closing Comments

The equations summarized in this Chapter will be used over and over again in the subsequent chapters. For the parts that are new to you make up simple problems with specific numbers for vectors and matrices and show that you get the same answers using matrix and indicial notations. We move on fairly quickly, each step is not particularly difficult, but the material accumulates with a fair amount of repetition.