Appendix KK

Vector Spaces and Matrices

A *vector space* is defined to be a set of elements called vectors for which vector addition and scalar multiplication is defined. Vector addition assigns to every pair of vectors, x and y, a sum, $x + y$, in such a way that

(1) addition is commutative, $x + y = y + x$,
(2) addition is associative, $x + (y + z) = (x + y) + z$,
(3) there is in the vector space a unique vector 0 (called the origin) such that $x + 0 = x$ for every vector x, and
(4) to every vector x in the space there corresponds a unique vector $-x$ such that $x + (-x) = 0$.

The multiplication of a scalar α times a vector x assigns to every pair, α and x, a vector αx in the vector space called the product of α and x. Scalar-vector multiplication is such that

(1) multiplication by scalars is associative, $\alpha(\beta x) = (\alpha\beta)x$,
(2) $1x = x$ for every vector x,
(3) multiplication by scalars is distributive with respect to vector addition, $\alpha(x + y) = \alpha x + \alpha y$, and
(4) multiplication by scalars is distributive with respect to scalar addition, $(\alpha + \beta)x = \alpha x + \beta x$.

An elementary example of a vector space is ordered sets of n numbers

$$x_n = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

with vector addition and scalar multiplication defined by the equations

$$x_n + y_n = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_n \\ \vdots \\ x_n + y_n \end{bmatrix}, \quad ax_n = \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix}.$$

If the components of the vectors in this space are complex numbers, the vector space is denoted C^n. The vector space consisting of ordered sets of real numbers is denoted R^n. Another example of a vector space is the set of polynomials

$$p_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

with complex coefficients a_n. The addition of two polynomial and multiplication by a complex number is defined in the ordinary way.

A finite set $\{x_i\}$ of vectors is said to be *linearly dependent* if there is a corresponding set $\{\alpha_i\}$ of numbers, not all zero, such that

$$\sum_i \alpha_i x_i = 0,$$

where the zero on the right-hand side of the equation corresponds to the zero vector. If, on the other hand, $\sum_i \alpha_i x_i = 0$ implies that $\alpha_i = 0$ for each i, the set $\{x_i\}$ is *linearly independent*. To illustrate the idea of linear dependence, we consider the following equation involving two vectors

$$\alpha_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Using the rules we have given previously for scalar multiplication and vector addition in C^n, we can combine the two vectors on the left-hand side of the equation to obtain

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This last equation can only be true if the coefficients, α_1 and α_2, are both equal to zero. We thus conclude that the vectors, $[10]^T$ and $[01]^T$, are linearly independent. A basis in a vector space ν is a set of linearly independent vectors χ such that every vector in the space can be expressed as a linear combination of members of χ. For instance, the vectors,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

form a basis in the space C^2.

An *inner product* in a vector space is a complex or real valued function of the ordered pair of vectors x and y such that

(1) $(x, y) = \overline{(y, x)}$, where the line over the ordered pair on the right indicates complex conjugation,
(2) $(x, \alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 (x, y_1) + \alpha_2 (x, y_2)$,
(3) $(x, x) \geq 0$; $(x, x) = 0$ if and only if $x = 0$.

The condition (1) implies that (x, x) is always real, so that the inequality in (3) makes sense. In a space with an inner product defined, the *norm* of a vector $||x||$ is defined

$$||x|| = \sqrt{(x, x)}.$$

The inner product thus makes it possible to associate a norm or length with every vector in the space. For the space C^n, the inner product of two vectors

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

is defined

$$(x, y) = \sum_{i=1}^{n} \bar{x}_i y_i.$$

Now for a little terminology. The vectors, x and y, are said to be *orthogonal* if the inner product of the two vectors is equal to zero

$$(x,\ y) = 0.$$

A vector x is said to be normalized if its norm is equal to one

$$||x|| = 1,$$

and a basis of vectors $\{\phi_i\}$ is said to be *orthonormal* if each basis vector is orthogonal to the other members of the basis and if each basis vector is normalized.

The wave functions representing the states of a physical system may be thought of as vectors in a function space. The inner product of two wave functions, ψ_1 and ψ_2, which depend upon a single variable x is defined

$$(\psi, \phi) = \int \overline{\psi_1(x)} \psi_2(x)\, dx,$$

and the inner product for wave functions depending on two or three variables is defined accordingly. For a particle moving in three dimensions, the inner product of the wave functions, ψ_1 and ψ_2, is

$$(\psi, \phi) = \int \overline{\psi_1(\mathbf{r})}\psi_2(\mathbf{r})\,\mathrm{d}V,$$

where $\mathrm{d}V$ is the volume element.

The presence of a basis in a vector space makes it possible to associate a column vector in C^n with every vector in the space and to associate a matrix with operators acting on vectors in the space. Let ν be a vector space and let $\chi = \phi_1,\ \phi_2, \ldots, \phi_n$ be a basis of χ. Using the basis, a vector x may be expressed

$$x = \sum_{i=1}^{n} x_i\phi_i, \tag{KK.1}$$

and we may associate a column vector,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

with each vector x.

The product of an operator A and a vector x is a vector which may also be expressed as a linear combination of the basis vectors ϕ_i. This will be true when A acts on the members of the basis itself

$$A\phi_j = \sum_{i=1}^{n} a_{ij}\phi_i, \tag{KK.2}$$

for $j = 1, \ldots, n$. The set $\{a_{ij}\}$ of numbers, indexed with the double subscript i, j is the *matrix* corresponding to A. We shall generally denote the matrix of A by $[A]$. The matrix may be written more explicitly in the form of a square array

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}. \tag{KK.3}$$

If the basis is orthonormal, an explicit expression can be derived for the matrix elements a_{ij} by using the inner product. Taking the inner product of Eq. (KK.2) from the left with ϕ_k, gives

$$(\phi_k,\ A\phi_j) = \left(\phi_k,\ \sum_{i=1}^{n} a_{ij}\phi_i\right) = \sum_{i=1}^{n} a_{ij}(\phi_k,\ \phi_i) = a_{kj}.$$

To derive this last equation, we have used the second property of the inner product and the fact that the basis is orthonormal. The last result may be written

$$a_{kj} = (\phi_k,\ A\phi_j).$$

The result of acting with an operator A upon a vector x can be obtained from the matrix associated with A and the column vector associated with x. Using Eqs. (KK.1) and (KK.2), we obtain

$$Ax = A\left(\sum_{i=1}^{n} x_i\phi_i\right) = \sum_{i=1}^{n} x_i A\phi_i = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ji}\phi_i = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ji}x_i\right)\phi_j.$$

We may write

$$(Ax)_j = \sum_{i=1}^{n} a_{ji}x_i.$$

The jth component of the vector Ax may thus be obtained by forming the sum of the elements of jth row of the matrix $[A]$ times the components of the column vector $[x]$. As an example of this rule, we give the result of the matrix-vector multiplications in Problem 10.7:

$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix},$$
$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix},$$
$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix}.$$

The matrix corresponding to the product of two operators, A and B, may be obtained by allowing to AB to act upon an element of the basis

$$(AB)\phi_j = A(B\phi_j) = A\left(\sum_{k=1}^{n} b_{kj}\phi_k\right) = \sum_{k=1}^{n} b_{kj}A\phi_k,$$
$$\sum_{k=1}^{n} b_{kj}\left(\sum_{i=1}^{n} a_{ik}\phi_i\right) = \sum_{i=1}^{n}\left(\sum_{i=1}^{n} a_{ik}b_{kj}\right)\phi_i.$$

We thus define the matrix product $[A][B]$ by the equation

$$([A][B])_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}. \tag{KK.4}$$

The process of forming the product of two matrices can be described in terms of the individual matrices. To obtain the ijth element of the product matrices, one forms the sum of the products of the elements of the ith row of $[A]$ with the elements of the jth column of $[B]$. The matrix multiplication will be well defined only if the matrix $[A]$ has as many columns as the matrix $[B]$ has rows. As an example of the rule (KK.4) for multiplying matrices, we give the result of the matrix-matrix multiplications in Problem 10.8:

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix},$$
$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix},$$
$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 3 \\ 5 & 3 & 2 \\ 6 & 2 & 4 \end{bmatrix}.$$