CHEM1034 - Week 31 Lecture - vectors and matrices

Sections 16.1-16.5, 16.10, 18.1-18.3 of Steiner, "The Chemistry Maths Book", 2nd edition.

Finite-dimensional vector spaces

A *vector* is defined as an *ordered set of numbers*. Those numbers (called *scalars*) may come from any number field; unless otherwise stated, we shall assume that they are complex. For historical reasons the numbers in a vector are usually written out as a column. Elementary operations that may be performed on complex vectors are:

addition
$$\vec{a}+\vec{b}=|a\rangle+|b\rangle=\begin{pmatrix}a_1\\a_2\\\vdots\\a_N\end{pmatrix}+\begin{pmatrix}b_1\\b_2\\\vdots\\b_N\end{pmatrix}=\begin{pmatrix}a_1+b_1\\a_2+b_2\\\vdots\\a_N+b_N\end{pmatrix}$$

multiplication by a scalar
$$c\vec{a}=c|a\rangle=c\begin{pmatrix}a_1\\a_2\\\vdots\\a_N\end{pmatrix}=\begin{pmatrix}ca_1\\ca_2\\\vdots\\ca_N\end{pmatrix}$$

Arrow notation ($\vec{a}$) or bold font ($\mathbf{a}$) are used in elementary mathematics and angle bracket notation $|a\rangle$ in quantum theory. The latter (proposed by Paul Dirac) is very convenient and is slowly taking over.

A set of vectors V over a field $\mathbb{F}$ is called a *vector space* if

1. The set is closed under addition and multiplication by a scalar:
$$\forall\vec{a},\vec{b}\in V\quad \vec{a}+\vec{b}\in V$$
$$\forall\vec{a}\in V\quad \forall\alpha\in\mathbb{F}\quad \alpha\vec{a}\in V$$
2. The addition operation is associative and commutative:
$$\forall\vec{a},\vec{b}\in V\quad \vec{a}+\vec{b}=\vec{b}+\vec{a}$$
$$\forall\vec{a},\vec{b},\vec{c}\in V\quad \left(\vec{a}+\vec{b}\right)+\vec{c}=\vec{a}+\left(\vec{b}+\vec{c}\right)$$
3. There exists a unique zero vector:
$$\exists!\vec{0}\in V,\quad \forall\vec{a}\in V\quad \vec{a}+\vec{0}=\vec{0}+\vec{a}=\vec{a}$$
4. There exists a unique additive inverse for each vector:
$$\forall\vec{a}\in V\quad \exists!\vec{b}\in V,\quad \vec{a}+\vec{b}=\vec{0}$$
5. Multiplication by scalars is compatible with field multiplication:
$$\forall\vec{a}\in V\quad \forall\alpha,\beta\in\mathbb{F}\quad \alpha\left(\beta\vec{a}\right)=\left(\alpha\beta\right)\vec{a}$$
6. Distributivity relations hold for field addition and vector addition:
$$\forall\vec{a},\vec{b}\in V\quad \forall\alpha\in\mathbb{F}\quad \alpha\left(\vec{a}+\vec{b}\right)=\alpha\vec{a}+\alpha\vec{b}$$
$$\forall\vec{a}\in V\quad \forall\alpha,\beta\in\mathbb{F}\quad \left(\alpha+\beta\right)\vec{a}=\alpha\vec{a}+\beta\vec{a}$$

The number N of elements in each vector in V is called the *dimension* of the vector space. There are subtle differences between finite- and infinite-dimensional vector spaces that we would not mention in this course – the number N will henceforth be assumed to be finite. All *N*-dimensional vector spaces over a given field have exactly the same properties. *N*-dimensional real and complex spaces are denoted $\mathbb{R}^N$ and $\mathbb{C}^N$ respectively. Exercise: demonstrate that $\mathbb{R}^3$ satisfies the definition of a vector space.

An operation that is often useful when working with vector spaces is *conjugate-transpose*:

$$\vec{a}^\dagger = \langle a| = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}^\dagger = \begin{pmatrix} a_1^* & a_2^* & \cdots & a_N^* \end{pmatrix}$$

It is performed by rewriting vectors in a row (a matter of convenience as we shall see below) and conjugating every element. Conjugate-transpose operation is denoted by a dagger or by inverting the direction of the bracket in the Dirac notation. In the case of real vector spaces the operation is called *transpose* because complex conjugation has no effect on real numbers.

Vector norm and scalar product

The generalization of the notion of vector length to spaces of arbitrary dimension is called the *norm*. Many different functions can serve as norms, but the most popular norm in physical sciences is

$$|\vec{a}| = \sqrt{\sum_{n=1}^{N} |a_n|^2} = \sqrt{\sum_{n=1}^{N} a_n^* a_n}$$

that is, the squares of the moduli of all elements should be added up and a square root taken. For $N=3$ this equation reduces to the standard expression for the length of a three-dimensional vector. A vector is called *normalized* if its norm is equal to 1.

The sum under the square root has two alternative notations:

$$\sum_{n=1}^{N} a_n^* a_n = \langle a|a\rangle = \begin{pmatrix} a_1^* & a_2^* & \cdots & a_N^* \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}$$

The scalar product of two vectors can also be generalized to complex spaces of arbitrary dimension. This is where Dirac notation becomes particularly useful:

$$\left(\vec{a} \cdot \vec{b}\right) = \sum_{n=1}^{N} a_n^* b_n = \langle a|b\rangle = \begin{pmatrix} a_1^* & a_2^* & \cdots & a_N^* \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}$$

The geometric meaning of the scalar product as an indicator of the angle between the two vectors

$$\cos \varphi_{\vec{a},\vec{b}} = \frac{\langle a|b\rangle}{|\vec{a}| \cdot |\vec{b}|}$$

is preserved in spaces of dimension higher than 3. If the scalar product of two non-zero vectors is zero, they are called *orthogonal*. A space of dimension *N* can have at most *N* mutually orthogonal vectors.

Matrices and linear maps

A *matrix* is a ordered array of numbers, usually written out as a table. A good practical example is a bitmap image, where values of red, green and blue pixel intensities are stored as matrices. The number of elements along its two dimensions is, in general, different. The elements are indexed as $a_{\text{row, column}}$:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,M} \\ a_{21} & a_{22} & \cdots & a_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,M} \end{pmatrix}$$

It can be easily demonstrated that the set of all matrices of a given dimension is also a vector space under addition and multiplication by a scalar – do it as an exercise. Matrices are interesing in their own right, but their primary function in physical sciences is to provide maps between vector spaces, which are accomplished by matrix-vector multiplication:

$$\mathbf{A}\vec{b} = \mathbf{A}|b\rangle = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,M} \\ a_{21} & a_{22} & \cdots & a_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,M} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{pmatrix} = \begin{pmatrix} \sum_{m=1}^{M} a_{1m} b_m \\ \sum_{m=1}^{M} a_{2m} b_m \\ \vdots \\ \sum_{m=1}^{M} a_{N,m} b_m \end{pmatrix} \quad \Leftrightarrow \quad \left[\mathbf{A}\vec{b}\right]_k = \sum_{m=1}^{M} a_{km} b_m$$

That is to say, the first row of the matrix is multiplied, element by element, with the vector and the result makes up the first element of the answer. Then the second row is multiplied, element by element, with the vector and the result makes up the second element of the answer. This process is continued until there are no further rows in the matrix. Clearly, the result is a vector of dimension *N*.

Matrix-vector multiplication produces another vector, generally of a different dimension. The multiplication by a matrix may therefore be viewed as an operation that creates a connection between different vector spaces. Such a connection is called a *linear map*.

Note that *the column dimension of the matrix must be the same as the row dimension of the vector*, otherwise the numbers of elements do not match. Examples:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1\cdot a + 2\cdot b + 3\cdot c \\ 4\cdot a + 5\cdot b + 6\cdot c \\ 7\cdot a + 8\cdot b + 9\cdot c \end{pmatrix}, \qquad \begin{pmatrix} 2+i & 1-4i & 3+2i \\ 5-i & -3+2i & 6+3i \\ 1+6i & 5i & 5+2i \end{pmatrix} \begin{pmatrix} 3 \\ 9+i \\ 1-i \end{pmatrix} = \begin{pmatrix} 24-33i \\ -5+9i \\ 5+60i \end{pmatrix}$$

Matrix-matrix multiplication is performed in a similar way:

$$\left[\mathbf{A}\cdot\mathbf{B}\right]_{nk} = \sum_m a_{nm} b_{mk}$$

In practice, a *row* of the left matrix is multiplied element-wise by a *column* of the right matrix, the result is summed up and placed into the corresponding *row and column* of the result. Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 4 \\ 10 & 5 & 10 \\ 16 & 8 & 16 \end{pmatrix}$$

Week 32 workshop exercises

Steiner, 2nd edition: section 16.11, problems 6-11, 19-26; section 18.8, problems 13, 17, 22, 26, 28.

Extra difficulty exercises for the brave

Steiner, 2nd edition: section 16.11, problems 16-18; section 18.8, problems 36, 37, 42.