CHEM2024 - Week 18 Lecture 1 - Vector and matrix spaces I

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

Systems of linear equations are ubiquitous in physical sciences. Their general form is:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \tag{1}$$

When the need to solve large systems of this kind appeared in the early 20th century, it quickly became clear that the best way to manipulate such expressions is to separate coefficient arrays from the array of variables. The following notation was adopted for Equation (1):

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad \Leftrightarrow \quad \mathbf{A}\vec{x} = \vec{b} \tag{2}$$

where $\vec{x}$ and $\vec{b}$ are vectors, and $\mathbf{A}$ is called a matrix. By about 1960-es it also became clear that this notation is very convenient for data processing using computers.

1. Vector spaces

A *vector* is defined as an *ordered set of numbers*. Those numbers (called *scalars*) may come from any number field; we shall assume that they are complex. Elementary operations that may be performed on complex vectors are:

addition

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

multiplication by a scalar

$$\alpha \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}$$

transpose

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

conjugate-transpose
$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}^{\dagger} = \begin{pmatrix} x_1^* & x_2^* & \cdots & x_n^* \end{pmatrix}$$

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:
$$\forall \vec{x},\vec{y} \in V \quad \vec{x}+\vec{y} \in V$$
$$\forall \vec{x} \in V \quad \forall \alpha \in \mathbb{F} \quad \alpha\vec{x} \in V$$
2. The addition operation is associative and commutative:
$$\forall \vec{x},\vec{y} \in V \quad \vec{x}+\vec{y} = \vec{y}+\vec{x}$$
$$\forall \vec{x},\vec{y},\vec{z} \in V \quad \left(\vec{x}+\vec{y}\right)+\vec{z} = \vec{x}+\left(\vec{y}+\vec{z}\right)$$
3. There exists a unique zero vector:
$$\exists ! \vec{0} \in V, \quad \forall \vec{x} \in V \quad \vec{x}+\vec{0} = \vec{0}+\vec{x} = \vec{x}$$
4. There exists a unique additive inverse for each vector:
$$\forall \vec{x} \in V \quad \exists ! \vec{y} \in V, \quad \vec{x}+\vec{y} = \vec{0}$$
5. Associativity relations hold for multiplication by scalars:
$$\forall \vec{x} \in V \quad \forall \alpha,\beta \in \mathbb{F} \quad \alpha\left(\beta\vec{x}\right) = \left(\alpha\beta\right)\vec{x}$$
6. Distributivity relations hold for addition and multiplication by scalars:
$$\forall \vec{x},\vec{y} \in V \quad \forall \alpha \in \mathbb{F} \quad \alpha\left(\vec{x}+\vec{y}\right) = \alpha\vec{x}+\alpha\vec{y}$$
$$\forall \vec{x} \in V \quad \forall \alpha,\beta \in \mathbb{F} \quad \left(\alpha+\beta\right)\vec{x} = \alpha\vec{x}+\beta\vec{x}$$

Example 1: demonstrate that the three-dimensional Euclidean space satisfies the definition of a vector space. This space is called $\mathbb{R}^3$.

Solution: by checking the six properties one by one, we can demonstrate that they are satisfied.

2. Linear combinations and linear independence

A vector $\vec{y} \in V$ is called a *linear combination* of vectors $\vec{x}_1,...,\vec{x}_n \in V$ if

$$\vec{y} = \alpha_1\vec{x}_1 + ... + \alpha_n\vec{x}_n \tag{3}$$

for some scalars $\alpha_1,...,\alpha_n$. The set of vectors $\left\{\vec{x}_1,...,\vec{x}_n\right\}$ is called *linearly independent* if none of them can be expressed as a linear combination of others. Otherwise, the set is linearly dependent.

Example 2: demonstrate that the following set of vectors is linearly dependent

$$\vec{x}_1 = \begin{pmatrix} 5 \\ 3 \end{pmatrix}, \quad \vec{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{x}_3 = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

Solution: by direct inspection, $\vec{x}_3 = \vec{x}_1 + 2\vec{x}_2$.

3. 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 the physical sciences is

$$\left|\vec{x}\right| = \sqrt{\sum_{k=1}^{n}\left|x_k\right|^2} = \sqrt{\sum_{k=1}^{n} x_k^* x_k} \tag{4}$$

where n is the number of elements in the vector. 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 1.

A measure of angle between vectors is called *scalar product*. Again, many definitions are possible, but the one used in physical sciences is:

$$\left(\vec{x}\cdot\vec{y}\right) = \sum_{k=1}^{n} x_n^* y_n \tag{5}$$

The interpretation of the scalar product is the same as it was in $\mathbb{R}^3$:

$$\left(\vec{x}\cdot\vec{y}\right) = \left|\vec{x}\right|\left|\vec{y}\right|\cos\varphi \tag{6}$$

where φ is the angle betwee the two vectors. 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.

It follows from Equation (5) that the norm is the root of the scalar product of the vector with itself:

$$\left|\vec{x}\right| = \sqrt{\left(\vec{x}\cdot\vec{x}\right)} \tag{7}$$

Because many applications of this formalism involve complex numbers, it is important to not forget the complex conjugate operation in Equations (4) and (5).

Example 3: normalize the following vectors

$$\vec{x} = \begin{pmatrix}1\\1\end{pmatrix}, \quad \vec{y} = \begin{pmatrix}1+i\\2\end{pmatrix}, \quad \vec{z} = \begin{pmatrix}-2\\0\\2\end{pmatrix}$$

Solution: we must calculate the norms and divide the vectors by those norms

$$\left|\vec{x}\right| = \sqrt{2} \qquad \left|\vec{y}\right| = \sqrt{6} \qquad \left|\vec{z}\right| = \sqrt{8}$$

therefore, the normalised vectors are:

$$\frac{\vec{x}}{\left|\vec{x}\right|} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \quad \frac{\vec{y}}{\left|\vec{y}\right|} = \frac{1}{\sqrt{6}}\begin{pmatrix}1+i\\2\end{pmatrix}, \quad \frac{\vec{z}}{\left|\vec{z}\right|} = \frac{1}{\sqrt{8}}\begin{pmatrix}-2\\0\\2\end{pmatrix}$$

4. Alternative notation systems

Arrow notation ($\vec{x}$) or bold font ($\mathbf{x}$) are used in mathematics. In quantum mechanics, a particularly convenient bracket notation system proposed by Paul Dirac is used. In that notation, the scalar product is denoted by an angular bracket:

$$\left(\vec{x}\cdot\vec{y}\right) = \langle x|y\rangle, \qquad \vec{x}^\dagger = \langle x|, \qquad \vec{x} = |x\rangle \tag{8}$$

The "bra" component of the bracket is identified with the conjugate-transposed vector, and the "ket" component with the vector itself:

$$\langle x | y \rangle = \begin{pmatrix} x_1^* & x_2^* & \cdots & x_n^* \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{k=1}^{n} x_k^* y_k \tag{9}$$

This notation is standard in computational chemistry and will be used extensively in this course.