CHEM2024 - Week 19 Lecture 1 - Matrix functions and equations

Chapters 27 and 28 of Monk and Munro, "Maths for Chemistry", 2nd edition.
Chapter 18 of Steiner, "The Chemistry Maths Book", 2nd edition.

1. Matrices and linear maps

A *matrix* is an 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 $a_{\text{row, column}}$:

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

The set of all matrices of a given dimension is a vector space under addition and multiplication by a scalar. Their primary function in physical sciences is to provide maps between vector spaces, which are accomplished by matrix-vector multiplication:

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

Matrix-vector multiplication produces another vector, generally of a different dimension. This creates a connection, called a *linear map*, between different vector spaces.

For the multiplication to be possible *the number of columns in the matrix must be the same as the number of rows in the vector*

$$\begin{pmatrix} \bullet & \bullet & \bullet \\ & & \\ & & \end{pmatrix}\begin{pmatrix} \bullet \\ \bullet \\ \bullet \end{pmatrix} = \begin{bmatrix} \bullet \\ \\ \\ \end{bmatrix}$$

otherwise the numbers of elements would not match. Examples:

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

Matrix-matrix multiplication is performed in a similar way:

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

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.

$$\begin{pmatrix} \bullet & \bullet & \bullet \\ & & \\ & & \end{pmatrix}\begin{pmatrix} \bullet \\ \bullet \\ \bullet \end{pmatrix}=\begin{pmatrix} \bullet \\ \\ \\ \end{pmatrix}$$

Example:

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

2. Matrix transpose and conjugate-transpose

Matrix *conjugate-transpose* operation reflects the positions of matrix elements relative to the diagonal and conjugates each element:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,M} \\ a_{21} & a_{22} & \cdots & a_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,M} \end{pmatrix}^{\dagger}=\begin{pmatrix} a_{11}^{*} & a_{21}^{*} & \cdots & a_{N,1}^{*} \\ a_{12}^{*} & a_{22}^{*} & \cdots & a_{N,2}^{*} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,M}^{*} & a_{2,M}^{*} & \cdots & a_{N,M}^{*} \end{pmatrix} \tag{1}$$

Conjugate transpose is denoted with a dagger symbol: $\mathbf{A}^{\dagger}$. In the case of real matrices the operation is called simply *transpose* and is denoted with a T symbol: $\mathbf{A}^{\mathrm{T}}$.

3. Matrix commutation

Matrix multiplication is not commutative, *i.e.* in general $\mathbf{AB} \neq \mathbf{BA}$. This property has deep consequences in quantum mechanics, where it leads to uncertainty relations. The function that returns the difference between $\mathbf{AB}$ and $\mathbf{BA}$ is called a *commutator* of $\mathbf{A}$ and $\mathbf{B}$:

$$\left[\mathbf{A},\mathbf{B}\right]=\mathbf{AB}-\mathbf{BA} \tag{2}$$

If the commutator happens to be zero, it is said that the two matrices *commute*.

4. Matrix functions in general

The operations that are defined for matrices are addition and multiplication. Using Taylor series, these two operations may be used to construct an arbitrary matrix function:

$$f\left(x\right)=\sum_{n=0}^{\infty}\frac{f^{(n)}\left(0\right)}{n!}x^{n} \qquad \Rightarrow \qquad f\left(\mathbf{A}\right)=\sum_{n=0}^{\infty}\frac{f^{(n)}\left(0\right)}{n!}\mathbf{A}^{n} \tag{3}$$

A particulary useful matrix function is the *matrix exponential*:

$$\exp\left(\mathbf{A}\right)=\sum_{n=0}^{\infty}\frac{\mathbf{A}^{n}}{n!} \tag{4}$$

Because infinite sums of matrix products are involved, matrix functions are usually calculated numerically using a computer. Another useful function is matrix *inverse* $\mathbf{A}^{-1}$, which is defined as

$$\mathbf{AA}^{-1}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{1}, \qquad \mathbf{1}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5}$$

For the matrix exponential and the inverse to exist, the matrix must be square. For square matrices, the exponential always exists, but the inverse might not exist, *e.g.* for a zero matrix.

5. Common types of matrices

Several specific types of matrices occur frequencly in physical sciences:

1. Symmetric matrix: $\mathbf{A}^{\mathrm{T}}=\mathbf{A}$. Example: diffusion operator in numerical hydrodynamics.
2. Hermitian matrix: $\mathbf{A}^{\dagger}=\mathbf{A}$. Example: Hamiltonians in quantum mechanics.
3. Orthogonal matrix: $\mathbf{A}^{\mathrm{T}}=\mathbf{A}^{-1}$. Example: rotation matrices.
4. Unitary matrix: $\mathbf{A}^{\dagger}=\mathbf{A}^{-1}$. Example: time evolution operators in quantum mechanics.
5. Traceless matrix: the *trace* (the sum of all diagonal elements) is zero. Example: spin operators.
6. Degenerate matrix: a matrix in which rows or columns are linearly dependent.

6. Matrix form of linear differential equations

Systems of linear differential equations may be cast into a matrix form. For example, the kinetic equations describing the following reaction chain

$$\mathrm{A}\underset{k_{1-}}{\overset{k_{1+}}{\rightleftarrows}}\mathrm{B}\underset{k_{2-}}{\overset{k_{2+}}{\rightleftarrows}}\mathrm{C} \quad \Rightarrow \quad \begin{cases} dA/dt=-k_{1+}A+k_{1-}B \\ dB/dt=+k_{1+}A-\left(k_{2+}+k_{1-}\right)B+k_{2-}C \\ dC/dt=+k_{2+}B-k_{2-}C \end{cases} \tag{6}$$

may be written in the matrix form as follows

$$\frac{d}{dt}\begin{pmatrix} A(t) \\ B(t) \\ C(t) \end{pmatrix}=\begin{pmatrix} -k_{1+} & +k_{1-} & 0 \\ +k_{1+} & -\left(k_{2+}+k_{1-}\right) & +k_{2-} \\ 0 & +k_{2+} & -k_{2-} \end{pmatrix}\begin{pmatrix} A(t) \\ B(t) \\ C(t) \end{pmatrix} \quad \Leftrightarrow \quad \frac{d}{dt}\vec{c}(t)=\mathbf{K}\vec{c}(t) \tag{7}$$

where $\mathbf{K}$ is called the *kinetic matrix* and $\vec{c}(t)$ is called the *concentration vector*. The law of the conservation of matter requires all column sums of $\mathbf{K}$ to be zero. The solution to Equation (7) is remarkably simple and may be written *via* a matrix exponential:

$$\frac{d}{dt}\vec{c}(t)=\mathbf{K}\vec{c}(t) \quad \Rightarrow \quad \vec{c}(t)=\exp(\mathbf{K}t)\vec{c}(0) \tag{8}$$

The proof for this relationship may be given using Taylor series in Equation (4). Solving Equations (6) using conventional ODE techniques would take a very long time, but the exponential solution is simple.

7. Rotation matrices

A particular class of geometric transformations that benefits from matrix notation is rotations. For a vector in two dimensions, we know that

$$\begin{cases} x'=+x\cos\varphi-y\sin\varphi \\ y'=+x\sin\varphi+y\cos\varphi \end{cases} \quad \Rightarrow \quad \begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \quad \Rightarrow \quad \vec{r}\,'=\mathbf{R}\vec{r} \tag{9}$$

The matrix $\mathbf{R}$ is called the *rotation matrix*. In three dimensions, rotations around X, Y and Z axes may be constructed from Equation (9) by rearranging its elements so that they act in the required plane:

$$\mathbf{R}_{\mathrm{X}}=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\varphi & -\sin\varphi\\ 0 & \sin\varphi & \cos\varphi\end{pmatrix},\quad \mathbf{R}_{\mathrm{Y}}=\begin{pmatrix}\cos\varphi & 0 & -\sin\varphi\\ 0 & 1 & 0\\ \sin\varphi & 0 & \cos\varphi\end{pmatrix},\quad \mathbf{R}_{\mathrm{Z}}=\begin{pmatrix}\cos\varphi & -\sin\varphi & 0\\ \sin\varphi & \cos\varphi & 0\\ 0 & 0 & 1\end{pmatrix} \tag{10}$$

In this notation, sequential rotations around multiple axes become particularly easy because they may be accomplished by matrix multiplication. For example, a rotation by an angle γ around the Z axis, followed by a rotation by an angle β around the Y axis, followed by another rotation by an angle α around the Z axis (the so-called *Euler angles*) may be written as:

$$\mathbf{R}(\alpha,\beta,\gamma)=\mathbf{R}_{\mathrm{Z}}(\alpha)\cdot\mathbf{R}_{\mathrm{Y}}(\beta)\cdot\mathbf{R}_{\mathrm{Z}}(\gamma) \tag{11}$$

note that the matrices are written from right to left. This is because when multiple matrices are multiplied into a vector, the nearest matrix acts first:

$$\mathbf{R}(\alpha,\beta,\gamma)\vec{r}=\mathbf{R}_{\mathrm{Z}}(\alpha)\cdot\mathbf{R}_{\mathrm{Y}}(\beta)\cdot\mathbf{R}_{\mathrm{Z}}(\gamma)\vec{r} \tag{12}$$

It is also easy to show that a rotation by a negative angle produces a matrix that is the inverse of the matrix that rotates around the same axis by a positive angle:

$$\mathbf{R}_{\mathrm{Z}}(+\varphi)\cdot\mathbf{R}_{\mathrm{Z}}(-\varphi)=\begin{pmatrix}\cos\varphi & -\sin\varphi & 0\\ \sin\varphi & \cos\varphi & 0\\ 0 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix}\cos\varphi & \sin\varphi & 0\\ -\sin\varphi & \cos\varphi & 0\\ 0 & 0 & 1\end{pmatrix}=\ldots=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \tag{13}$$

Rotation matrices are an example of an important mathematical concept called *matrix representation* – every rotation may be uniquely associated with a matrix and the behaviour of those matrices under multiplication *identically repeats* the behaviour of rotations under superposition. Many types of physical operators have matrix representations. This simplifies the practical mathematics because matrices are easy to multiply on a computer.

Rotation matrices in three dimensions are also an example of a *non-commutative group* – the order in which rotations are applied does matter. For example:

$$\mathbf{R}_{\mathrm{X}}(\alpha)\cdot\mathbf{R}_{\mathrm{Y}}(\beta)\neq\mathbf{R}_{\mathrm{Y}}(\beta)\cdot\mathbf{R}_{\mathrm{X}}(\alpha) \tag{14}$$

However, two-dimensional rotations (and rotations around the same axis in general) do commute, *e.g.*:

$$\mathbf{R}_{\mathrm{Z}}(\alpha)\cdot\mathbf{R}_{\mathrm{Z}}(\beta)=\mathbf{R}_{\mathrm{Z}}(\beta)\cdot\mathbf{R}_{\mathrm{Z}}(\alpha) \tag{15}$$

The proof of equations (14) and (15) is left as a homework exercise.