CHEM1030 - Week 23 Lecture - 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.

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: $\mathbf{A}^{\mathrm{T}}$.

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*.

Matrix functions in general

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

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

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

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

which inherits the infinite convergence radius from the original Taylor series. 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

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

Note that, for the matrix exponential and the inverse to exist, the matrix must be square. For square matrices, the exponential always exists, but the inverse may or may not exist, depending on the matrix.

Matrix form of linear algebraic equations

For two vectors to be equal, they must be equal *element by element*. This provides a connection between vector equalities and systems of equations, for example:

$$\vec{a}=\vec{b} \quad \Leftrightarrow \quad \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} a_1 = b_1 \\ a_2 = b_2 \\ \vdots \\ a_N = b_N \end{cases} \tag{6}$$

In particular, this allows one to recast systems of linear equations in matrix form:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1,N}x_N = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2,N}x_N = b_2 \\ \cdots \\ a_{N,1}x_1 + a_{N,2}x_2 + \cdots + a_{N,N}x_N = b_N \end{cases} \quad \Leftrightarrow \quad \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,N} \\ a_{21} & a_{22} & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} \tag{7}$$

This system of equations may therefore be written very compactly and also (assuming a unique solution exists) easily solved by multiplying both sides by $\mathbf{A}^{-1}$ and observing that $\mathbf{1}\vec{x}=\vec{x}$:

$$\mathbf{A}\vec{x}=\vec{b} \quad \Rightarrow \quad \mathbf{A}^{-1}\mathbf{A}\vec{x}=\mathbf{A}^{-1}\vec{b} \quad \Rightarrow \quad \mathbf{1}\vec{x}=\mathbf{A}^{-1}\vec{b} \quad \Rightarrow \quad \vec{x}=\mathbf{A}^{-1}\vec{b} \tag{8}$$

Systems with millions of equations are solved on modern computers in seconds using this method.

Matrix form of linear differential equations

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

$$\mathrm{A} \underset{k_{1-}}{\overset{k_{1+}}{\rightleftarrows}} \mathrm{B} \underset{k_{2-}}{\overset{k_{2+}}{\rightleftarrows}} \mathrm{C} \quad \Rightarrow \quad \begin{cases} dA/dt = -k_{1+}A + k_{1-}B \\ dB/dt = +k_{1+}A - \left(k_{2+}+k_{1-}\right)B + k_{2-}C \\ dC/dt = +k_{2+}B - k_{2-}C \end{cases} \tag{9}$$

may be written in the matrix form as follows

$$\frac{d}{dt}\begin{pmatrix} A(t) \\ B(t) \\ C(t) \end{pmatrix} = \begin{pmatrix} -k_{1+} & +k_{1-} & 0 \\ +k_{1+} & -\left(k_{2+}+k_{1-}\right) & +k_{2-} \\ 0 & +k_{2+} & -k_{2-} \end{pmatrix} \begin{pmatrix} A(t) \\ B(t) \\ C(t) \end{pmatrix} \quad \Leftrightarrow \quad \frac{d}{dt}\vec{c}(t) = \mathbf{K}\vec{c}(t) \tag{10}$$

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 (10) is remarkably simple and may be written *via* a matrix exponential:

$$\frac{d}{dt}\vec{c}(t) = \mathbf{K}\vec{c}(t) \quad \Rightarrow \quad \vec{c}(t) = \exp(\mathbf{K}t)\vec{c}(0) \tag{11}$$

The proof for this relationship may be given using Taylor series in Equation (4) and is left as an exercise. Note that solving Equations (9) using conventional ODE techniques would take a very long time.

Rotation matrices

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

$$\begin{cases} x' = +x\cos\varphi - y\sin\varphi \\ y' = +x\sin\varphi + y\cos\varphi \end{cases} \quad \Rightarrow \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \Rightarrow \quad \vec{r}\,' = \mathbf{R}\vec{r} \tag{12}$$

The matrix $\mathbf{R}$ is called the *rotation matrix*. In three dimensions, rotations around X, Y and Z axes may be constructed from Equation (12) by rearranging its elements so that they act in the required plane:

$$\mathbf{R}_X = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}, \quad \mathbf{R}_Y = \begin{pmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{pmatrix}, \quad \mathbf{R}_Z = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{13}$$

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:

$$\mathbf{R}(\alpha,\beta,\gamma) = \mathbf{R}_Z(\alpha)\cdot\mathbf{R}_Y(\beta)\cdot\mathbf{R}_Z(\gamma) \tag{14}$$

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:

$$\mathbf{R}(\alpha,\beta,\gamma)\vec{r} = \mathbf{R}_Z(\alpha)\cdot\mathbf{R}_Y(\beta)\cdot\mathbf{R}_Z(\gamma)\vec{r} \tag{15}$$

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:

$$\mathbf{R}_Z(+\varphi)\cdot\mathbf{R}_Z(-\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}\cdot\begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \ldots = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{16}$$

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:

$$\mathbf{R}_X(\alpha)\cdot\mathbf{R}_Y(\beta) \neq \mathbf{R}_Y(\beta)\cdot\mathbf{R}_X(\alpha) \tag{17}$$

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

$$\mathbf{R}_Z(\alpha)\cdot\mathbf{R}_Z(\beta) = \mathbf{R}_Z(\beta)\cdot\mathbf{R}_Z(\alpha) \tag{18}$$

The proof of equations (17) and (18) is left as a homework exercise.

Week 24 workshop exercises

Monk and Munro, 2nd edition: self-tests 27.3, 27.5, 27.6; additional problem 27.4

Steiner, 2nd edition: section 18.8, problems 54, 55, 60.

Extra difficulty exercise for the brave

Perform the multiplications in Equation (14) and write the resulting rotation matrix out explicitly. Show that for $\beta = 0$ the angles α and γ are not defined uniquely (several satellites have been lost by various space agencies due to this particular problem).