CHMC Advanced Group:
To Quaternions and beyond!

3/23/2017

1 Introduction

An important theme in mathematics is encoding certain objects and the structure of interactions between these objects in a variety of different representations. By viewing the same objects in different contexts, one may learn about the structure that is associated with these objects; often, this structure may be obscured from certain points of view and more clear in other points of view. One common way to encode objects and their structure is through the use of matrices, about which much is known and well-understood. In this worksheet, we explore how to represent different number systems in the form of matrices and how this representation actually allows us to generalize this approach into new number systems.

2 Basics on Matrices and Matrix Operations

In this section we'll introduce the concept of a matrix, as well as how to work with them. This section is just a collection of tools and won't seem very motivated. Don't worry though, we'll see plenty of cool ways that matrices are utilized in the sections that follow.

2.1 What is a matrix?

The most basic interpretation of a matrix is a rectangular array of numbers. For example, these are all matrices:

$$\begin{pmatrix} 3 & 7 \\ 0.5 & 2 \end{pmatrix}, \begin{pmatrix} 2 & \pi & 2\pi \\ \sqrt{2} & \frac{1}{\pi} & 300 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 \\ 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 \end{pmatrix}, \begin{pmatrix} 1.12345 \end{pmatrix}.$$

There isn't any extra structure we put on these arrays, they are just organized collections of numbers. In what follows, we'll usually refer to the numbers of the array as *elements* of the matrix. The first matrix above is a 2×2 matrix, since it has two rows and two columns. The second is a 2×3 matrix, since it has two rows and three columns.

Exercise 2.1 What kinds of matrices are the last two above? In other words, they are $m \times n$ matrices where m and n are integers; what are the integers?

In many situations, we think of 1×1 matrices as a real number, since they are more-or-less the same kind of object.

2.2 How do we add matrices?

Now we'll put a little more structure on these objects, starting with addition. For two matrices of the same type (say, two 2×2 matrices, or two 3×4 matrices, etc.), we define addition element-wise. These examples should make this clear:

$$\begin{pmatrix} 2 & 5 & 8 \\ 1 & \pi & 0 \end{pmatrix} + \begin{pmatrix} 1 & 5 & -2 \\ 0 & 3\pi & 1 \end{pmatrix} = \begin{pmatrix} 2+1 & 5+5 & 8+(-2) \\ 1+0 & \pi+3\pi & 0+1 \end{pmatrix} = \begin{pmatrix} 3 & 10 & 6 \\ 1 & 4\pi & 1 \end{pmatrix},$$
$$\begin{pmatrix} 3 & 4 \\ 2 & -7 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 18 \end{pmatrix} = \begin{pmatrix} 3+1 & 4+2 \\ 2+0 & -7+18 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 2 & 11 \end{pmatrix}.$$

This is why we need the matrices to be of the same type, since otherwise one of the elements might not have a corresponding element to be added to.

Exercise 2.2 What matrix do you get from adding $\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$?

Exercise 2.3 What is the additive inverse of the matrix $\begin{pmatrix} 7 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix}$? In other words, find the matrix $\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$ such that

$$\begin{pmatrix} 7 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix} + \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Exercise 2.4 Using the last exercise (or otherwise), find the additive inverse of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where a, b, c, d can be any real number.

We call the matrix filled with 0s the zero matrix (who would've thought), and write it as just 0. For example, the 2×2 zero matrix is $0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Another important matrix (especially in the upcoming sections) is the *identity matrix*, which we'll write as I. This is the $m \times m$ matrix with 1s along the diagonal and 0s elsewhere. In the 2×2 case, it looks like $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

2.3 How do we multiply matrices?

Now we'll introduce an even fancier structure to our collection of matrices, multiplication. If A is an $m \times n$ matrix and B is an $n \times o$ matrix, then we define their *product* in a slightly strange way, which I'll show with some examples. It's important to note that the number of columns of A has to be the same as the number of rows of B!

Explicitly for 2×2 matrices (which we'll mainly be using):

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 8 & 0 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 1 \cdot 8 + 2 \cdot (-1) & 1 \cdot 0 + 2 \cdot 4 \\ 3 \cdot 8 + 2 \cdot (-1) & 3 \cdot 0 + 2 \cdot 4 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 22 & 8 \end{pmatrix},$$

$$\text{and so in general } \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \cdot \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = \begin{pmatrix} a_1 \cdot b_1 + a_2 \cdot b_3 & a_1 \cdot b_2 + a_2 \cdot b_4 \\ a_3 \cdot b_1 + a_4 \cdot b_3 & a_3 \cdot b_2 + a_4 \cdot b_4 \end{pmatrix}.$$

For other types of matrices, multiplication takes the same kind of form; we match elements in the rows of the first matrix with elements in the column of the second matrix, then add them up. Another type of matrix multiplication that we'll make extensive use of is the multiplication of a 2×2 matrix by a 2×1 matrix:

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \cdot (-2) + 2 \cdot 4 \\ 3 \cdot (-2) + 2 \cdot 4 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix},$$

$$\text{and so in general } \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 \cdot b_1 + a_2 \cdot b_2 \\ a_3 \cdot b_1 + a_4 \cdot b_2 \end{pmatrix}.$$

Since we don't need to multiply other types of matrices, I won't go into the general case, but here's an example of a 2×2 matrix multiplied with a 2×3 matrix (notice that the dimensions match up in the way they should!):

$$\begin{pmatrix}1 & 2\\ 3 & 2\end{pmatrix}\cdot\begin{pmatrix}8 & 0 & 1\\ -1 & 4 & 3\end{pmatrix} = \begin{pmatrix}1\cdot 8+2\cdot(-1) & 1\cdot 0+2\cdot 4 & 1\cdot 1+2\cdot 3\\ 3\cdot 8+2\cdot(-1) & 3\cdot 0+2\cdot 4 & 3\cdot 1+2\cdot 3\end{pmatrix} = \begin{pmatrix}6 & 8 & 7\\ 22 & 8 & 9\end{pmatrix}.$$

Exercise 2.5 Compute $\begin{pmatrix}3 & 2\\ 1 & 2\end{pmatrix}\cdot\begin{pmatrix}0 & 2\\ 3 & 0\end{pmatrix}$. What is $\begin{pmatrix}3 & 2\\ 1 & 2\end{pmatrix}\cdot\begin{pmatrix}0\\ 3\end{pmatrix}$? Notice any connections between these two matrix products?

Exercise 2.6 Compute $\begin{pmatrix}3 & 2\\ 1 & 2\end{pmatrix}\cdot I = \begin{pmatrix}3 & 2\\ 1 & 2\end{pmatrix}\cdot\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$. What do you notice?

The last exercise shows why we call I the identity matrix; it's the multiplicative identity for matrices (just like how 1 is the multiplicative identity for real numbers)!

Exercise 2.7 Compute $\begin{pmatrix}6 & 3\\ 2 & 1\end{pmatrix}\cdot\begin{pmatrix}1 & -2\\ -2 & 4\end{pmatrix}$. What do you notice?

Exercise 2.8 What about $\begin{pmatrix}1 & -2\\ -2 & 4\end{pmatrix}\cdot\begin{pmatrix}6 & 3\\ 2 & 1\end{pmatrix}$? What is the difference between matrix multiplication and ordinary real number multiplication?

Exercise 2.9 Given two 2×2 matrices A, B, what conditions do you need for $A\cdot B = B\cdot A$? If you're not sure, try a simpler case: what if $A = \begin{pmatrix}a_1 & a_2\\ a_3 & a_4\end{pmatrix}$ and $B = \begin{pmatrix}b_1 & 0\\ 0 & b_2\end{pmatrix}$? Do these commute?

In one of the last exercises you should have gotten the 0 matrix. Don't be alarmed! There are other places in math where multiplying two non-zero things gives zero, and matrix multiplication is one of those places. So long as we don't divide by zero, we won't run into any trouble.

Finally, we have another kind of multiplication for matrices. This time though we multiply the matrix by a real number, not another matrix, and the multiplication is done element-wise (just like addition). For example,

$$3\cdot\begin{pmatrix}2 & -2\\ -2 & 4\end{pmatrix} = \begin{pmatrix}3\cdot 2 & 3\cdot(-2)\\ 3\cdot(-2) & 3\cdot 4\end{pmatrix} = \begin{pmatrix}6 & -6\\ -6 & 12\end{pmatrix}.$$

2.4 The inverse of a 2×2 matrix

Now that we know how to add, subtract, and multiply matrices together, let's look into division.

When we divide two real numbers, say x/y, what we're really doing is looking for a number z such that $x = yz$. For example, $4/2$ (four divided by two) is equal to 2, because $4 = 2\cdot 2$. A good starting point is to find the inverse of numbers, like the multiplicative inverse of 2, and define division as the multiplication by the inverse. This is the approach we'll take.

To this end, we need to know what the inverse of a 2×2 matrix is: if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where $ad - bc \neq 0$, then the **inverse** of A is

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Exercise 2.10 Check that if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, with $ad - bc \neq 0$, then we do indeed have

$$A \cdot A^{-1} = I.$$

Thus, if we want to divide the matrix A by the matrix B, we multiply $A \cdot B^{-1}$. The only real number we can't divide by is 0, and this is encapsulated by the requirement that $\det(A) = ad - bc \neq 0$. Whenever you want to check whether something is invertible, compute the **determinant**[1] $\det(A)$ and see if that's 0.

Exercise 2.11 Is I invertible? Why or why not? What about 0?

Exercise 2.12 Is $\begin{pmatrix} 6 & 3 \\ 2 & 1 \end{pmatrix}$ invertible? Why or why not?

Look back at exercise 2.7, where the last matrix first appeared. It isn't a coincidence that the product there was 0, and that the matrix isn't invertible. In fact, if we only look at matrices that satisfy $\det(A) = ad - bc \neq 0$, i.e. the set

$$GL_2 = \{2 \times 2 \text{ matrices } A \text{ that satisfy } \det(A) \neq 0\},$$

then with matrix multiplication we get a group! Neat! In the following sections, it turns out to be this group structure that gives us all of the cool properties that we'll explore.

Exercise 2.13 * Let $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$. Let A' be the matrix we get by switching the columns of A. Show that $\det(A) = -\det(A')$.

Exercise 2.14 * Suppose that one of the columns of A is a multiple of the other, say $\begin{pmatrix} a_1 \\ a_3 \end{pmatrix} = c \begin{pmatrix} a_2 \\ a_4 \end{pmatrix}$ for some constant c. Show that in this case $\det(A) = 0$.

Exercise 2.15 * Show that, if A, B are 2×2 matrices, then $\det(A \cdot B) = \det(A) \cdot \det(B)$. Is $\det(A \cdot B) = \det(B \cdot A)$?

[1] The determinant is a funny function that may seem completely out of the blue and unmotivated. At this stage you're not wrong, but there are some deeper, more natural reasons as to why the determinant is defined the way it is (and it's an incredibly important function when working with matrices).

3 How do matrices act on points?

Now, we'll start to treat matrices as *maps*, instead of collections of numbers. This interpretation of matrices is central to the next two sections.

Exercise 3.1 What is $F \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$? Plot $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $F \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ on the same graph. Now plot the points $F^2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $F^3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, etc. What do you notice? Where are the points going?

Exercise 3.2 Do the same as in the last exercise, this time with the point $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Where do these points head off to?

We may view the length two column vectors as ordered 2-tuples in Euclidean space; in the above exercise, you interpreted the column vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ as the point $(1, 0)$ in $\mathbb{R}^2$. In general, for a column vector of length n, we may interpret this as a point in $\mathbb{R}^n$.

Recall that the complex numbers are of the form $z = a + bi$ where $a, b \in \mathbb{R}$ and i is defined to be the square root of -1. We plot the complex numbers $\mathbb{C}$ in the complex plane with horizontal axis corresponding to the real component, a in z above, and the vertical axis corresponding to the imaginary component, b in z above. In this way, we can also encode $\mathbb{C}$ as ordered pairs or 2-tuples of real numbers, points given by (a, b). The question is, can we encode more than just points in this way?

We will determine how to represent a complex number as a matrix of real numbers. To do this, we first examine what happens to a complex number $z = a + bi$ when multiplied by a real number c.

Exercise 3.3 Let $\begin{pmatrix} s & t \\ u & v \end{pmatrix}$ represent the real number c. What are the real and imaginary components of cz? Representing z as a column $\begin{pmatrix} a \\ b \end{pmatrix}$, what is the corresponding column vector for cz? What is the corresponding column vector for $\begin{pmatrix} s & t \\ u & v \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix}$ Set this equal to the column vector for cz. What are the values of s, t, u, v?

Exercise 3.4 Let $\begin{pmatrix} s & t \\ u & v \end{pmatrix}$ represent the purely imaginary number di. What are the real and imaginary components of diz? Representing z as a column $\begin{pmatrix} a \\ b \end{pmatrix}$, what is the corresponding column vector for diz? What is the corresponding column vector for $\begin{pmatrix} s & t \\ u & v \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix}$ Set this equal to the column vector for diz. What are the values of s, t, u, v?

Exercise 3.5 Putting together the last two exercises, what is the matrix the represents the complex number $w = c + di$?

Now, consider the complex numbers $z = a + bi$ and $w = c + di$.

Exercise 3.6 What are the matrix representations for z and w? What is the matrix representation of $(z + w)$? Compare this with adding the matrix representations of z and w. Also, compute the matrix representation of $(w + z)$. What is the matrix representation of zw? Compare this to the multiplying the matrix representations of z and w? How does this compare to the matrix representation of wz?

By representating the complex numbers as 2×2 matrices of real numbers, we have also encoded the structure given by addition and multiplication.

Exercise 3.7 Determine the conditions on a and b for which the matrix representing the complex number $z = a + bi$ has determinant 0. Recall that $a, b \in \mathbb{R}$.

Exercise 3.8 Let $z = a + bi$ be a non-zero complex number (so at least one of a, b is non-zero). Write this down in its matrix representation. What is the inverse of this matrix? What complex number does this correspond to? What happens when you multiply this complex number by z?

4 Quaternions

A **quaternion** q is a number of the form $q = a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}$, where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ satisfy the following properties:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1,$$
$$\mathbf{i}\,\mathbf{j} = \mathbf{k}, \mathbf{j}\,\mathbf{k} = \mathbf{i}, \text{and } \mathbf{k}\,\mathbf{i} = \mathbf{j}\,.$$

Exercise 4.1 Show, using the properties above, that $\mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1$.

Exercise 4.2 Show, using the properties above, that $\mathbf{j}\,\mathbf{i} = -\mathbf{k}, \mathbf{k}\,\mathbf{j} = -\mathbf{i}$, and $\mathbf{i}\,\mathbf{k} = -\mathbf{j}$.

Exercise 4.3 What is $\mathbf{i}^{-1}$? In otherwords, what quaternion q must we multiply $\mathbf{i}$ by to get $q \cdot \mathbf{i} = 1$?

Exercise 4.4 Show that $\mathbf{i}^{-1}\,\mathbf{j}\,\mathbf{i} = \mathbf{j}^{-1}$.

Quaternions have a multiplicative structure that behaves a lot like the complex and real numbers. For example, to multiply $p = a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}$ with $q = e + f\,\mathbf{i} + g\,\mathbf{j} + h\,\mathbf{k}$, we distribute the terms and then recombine:

$$\begin{aligned}
pq =& (a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k})(e + f\,\mathbf{i} + g\,\mathbf{j} + h\,\mathbf{k}) \\
=& ae + af\,\mathbf{i} + ag\,\mathbf{j} + ah\,\mathbf{k} + be\,\mathbf{i} + bf\,\mathbf{i}\,\mathbf{i} + bg\,\mathbf{i}\,\mathbf{j} + bh\,\mathbf{i}\,\mathbf{k} \\
&+ ce\,\mathbf{j} + cf\,\mathbf{j}\,\mathbf{i} + cg\,\mathbf{j}\,\mathbf{j} + ch\,\mathbf{j}\,\mathbf{k} + de\,\mathbf{k} + df\,\mathbf{k}\,\mathbf{i} + dg\,\mathbf{k}\,\mathbf{j} + dh\,\mathbf{k}\,\mathbf{k} \\
=& ae + af\,\mathbf{i} + ag\,\mathbf{j} + ah\,\mathbf{k} + be\,\mathbf{i} + bf(-1) + bg\,\mathbf{k} + bh(-\mathbf{j}) \\
&+ ce\,\mathbf{j} + cf(-\mathbf{k}) + cg(-1) + ch(\mathbf{i}) + de\,\mathbf{k} + df(\mathbf{j}) + dg(-\mathbf{i}) + dh(-1) \\
=& (ae - bf - cg - dh) + (af + be + ch - dg)\,\mathbf{i} \\
&+ (ag - bh + ce + df)\,\mathbf{j} + (ah + bg - cf + de)\,\mathbf{k}\,.
\end{aligned}$$

It is not recommended at all that you memorize this formula. Rather, whenever you need to multiply two quaternions go ahead and distribute the terms and do everything from scratch. Many of the quaternions we'll be working with in the last section will only have three non-zero components, rather than all four.

Exercise 4.5 Let $p = 1 + 3\,\mathbf{i} + \mathbf{k}$, $q = 2\,\mathbf{i} - 2\,\mathbf{j} + 3\,\mathbf{k}$, and $r = 4 + 3\,\mathbf{i} + 2\,\mathbf{j} + \mathbf{k}$. Compute pq, pr and rp.

Exercise 4.6 In general, suppose p and q are two quaternions. Is it true that $pq = qp$? Why or why not? If not, give an explicit example.

Exercise 4.7 If p, q, and r are three quaternions, is it true that $(pq)r = p(qr)$? Here, the parenthesis indicate which quaternions are to be multiplied together first. If so, show this.

Exercise 4.8 Complex numbers are a special case of quaternions: why?

We define the **magnitude** or **norm** of a quaternion $q = a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}$ to be the number

$$\|q\| = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

A useful definition, similar to that in the complex case, is that of a **conjugate** quaternion: if $q = a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}$, then the conjugate of q is the quaternion

$$\bar{q} = a - b\,\mathbf{i} - c\,\mathbf{j} - d\,\mathbf{k}\,.$$

Exercise 4.9 Show that $\|q\| = \sqrt{q\bar{q}}$.

The next exercise establishes a useful property that the quaternion norm satisfies, much like the ordinary absolute value does for real numbers.

Exercise 4.10 Show that if q and r are two quaternions, then $\|qr\| = \|q\|\|r\|$.

Exercise 4.11 Is it possible for $\|q\| = 0$, but $q \neq 0$? Why or why not? Recall that $q = 0$ means $a = b = c = d = 0$ for $q = a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}$.

In the last section we'll use the inverse of arbitrary quaternions, so let's verify that they take a particularly convenient form:

Exercise 4.12 Show that if $q = a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}$, then $q^{-1} = \frac{\bar{q}}{\|q\|^2} = \frac{a - b\,\mathbf{i} - c\,\mathbf{j} - d\,\mathbf{k}}{\|q\|^2}$.

Exercise 4.13 What is r^{-1}, where $r = 4 + 3\,\mathbf{i} + 2\,\mathbf{j} + \mathbf{k}$?

5 Representations of quaternions

In the first section, we noticed that the matrix $[\sqrt{-1}] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ satisfies the property that

$$[\sqrt{-1}]^2 = -I = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

which is the matrix analogue of the property that $\sqrt{-1}$ has. In this section we'll explore two representations of quaternions with matrices.

The first representation will involve matrices whose entries are complex numbers. Consider the two matrices

$$[\mathbf{i}]_1 = \begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix}, \quad [\mathbf{j}]_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Again, the subscript of 1 indicates that this is the first of two representations we'll consider in this section.

Exercise 5.1 Verify that, indeed, $([\mathbf{i}]_1)^2 = -I$ and $([\mathbf{j}]_1)^2 = -I$.

Exercise 5.2 Using the properties that $\mathbf{i}, \mathbf{j}$, and $\mathbf{k}$ must satisfy from before, what must the matrix $[\mathbf{k}]_1$ be? Verify that, for this matrix, we have $([\mathbf{k}]_1)^2 = -I$.

Exercise 5.3 Recall that $\mathbf{i}^{-1} = -\mathbf{i}$. What is $([\mathbf{i}]_1)^{-1}$? Use this, and the fact that $\mathbf{i}^{-1}\,\mathbf{j}\,\mathbf{i} = \mathbf{j}^{-1}$ to compute the matrix $([\mathbf{j}]_1)^{-1}$.

Now we look at writing an arbitrary quaternion $a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}$ as a matrix. The idea is to replace each of $\mathbf{i}, \mathbf{j}$, and $\mathbf{k}$ with their matrix equivalents. What do we do about the a term though? The most likely candidate is to replace the constant 1 with the matrix I, i.e. $[1]_1 = I$. In this way, we get

$$[a]_1 = a[1]_1 = aI = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

Thus, as a matrix, we have

$$[a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}]_1 = a[1]_1 + b[\mathbf{i}]_1 + c[\mathbf{j}]_1 + d[\mathbf{k}]_1.$$

The terms of the right hand side can be added together to obtain a single 2×2 matrix, giving the matrix representation of our quaternion.

Exercise 5.4 Let $q = 2\,\mathbf{i} - 2\,\mathbf{j} + 3\,\mathbf{k}$ and $r = 4 + 3\,\mathbf{i} + 2\,\mathbf{j} + \mathbf{k}$. What are the matrices $[q]_1$ and $[r]_1$? What is the matrix product $[q]_1[r]_1$? Does this agree with your findings from exercise 3.5?

Exercise 5.5 Let $q = a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}$. What is the matrix representation of q? What is $\bar{q}$ and its associated matrix? Compute the product of the matrices associated to $q\bar{q}$. What is the determinant? Relate this to the norm of q.

The second representation we'll consider will not involve complex numbers as matrix elements, but will involve 4×4 matrices with real entries. In a manner similar to before, let's set

$$[\mathbf{i}]_2 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad [\mathbf{j}]_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

Some of the intuition for why these matrices are chosen comes from the fact that $\mathbf{i} \cdot 1 = \mathbf{i}, \mathbf{i} \cdot \mathbf{i} = -1, \mathbf{i} \cdot \mathbf{j} = \mathbf{k}$, and $\mathbf{i} \cdot \mathbf{k} = -\mathbf{j}$. The matrix $[\mathbf{i}]_2$ keeps track of what happens when we multiply each of these "basic" quaternions $1, \mathbf{i}, \mathbf{j}$, and $\mathbf{k}$ by the quaternion $\mathbf{i}$. For example, the third column has a -1 in the fourth row, because multiplying the third "basic" quaternion $\mathbf{j}$ on the left by the quaternion $\mathbf{i}$ results in negative the fourth "basic" quaternion $\mathbf{k}$.

Exercise 5.6 With the same reasoning as in the last paragraph, why would we expect $[\mathbf{j}]_2$ to have the form that it does?

Exercise 5.7 Verify that

$$([\mathbf{i}]_2)^2 = ([\mathbf{j}])^2 = -I = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Exercise 5.8 Since $\mathbf{i}\,\mathbf{j} = \mathbf{k}$, what must $[\mathbf{k}]_2$ be?

Exercise 5.9 Let $q = 2\,\mathbf{i} - 2\,\mathbf{j} + 3\,\mathbf{k}$ and $r = 4 + 3\,\mathbf{i} + 2\,\mathbf{j} + \mathbf{k}$. What are the matrices $[q]_2$ and $[r]_2$? Check that the matrix product $[q]_2[r]_2$ agrees with what you found in exercises 3.5 and 4.4.