Supplementary material to:

Primer 2: Linear Algebra

From:

A Biologist's Guide to Mathematical Modeling in Ecology and Evolution

S. P. Otto and T. Day (2005)

Princeton University Press

Supplementary Material P2.1: More on matrix multiplication

When you see two matrices written as **A B**, this denotes the type of multiplication that we have been using. This operation is sometimes called the "*dot product*" or "*inner product*" and can also be written as $\mathbf{A} \bullet \mathbf{B} = \mathbf{A}\,\mathbf{B}$. There are, however, a few different multiplication operators that are occasionally seen in mathematical biology. These alternative products represent different operations that can be performed on lists of objects. Although we won't be using them in this book, it is worth knowing that they exist. The simplest multiplication operator is the "*Hadamard product*" or "*direct product*" of two $r \times c$ matrices, written as $\mathbf{A} \circ \mathbf{B}$. The Hadamard product is straightforward to calculate: you just multiply the two matrices together element by element, placing A_{ij} times B_{ij} in the ij^{th} position of the new $r \times c$ matrix. For example, if you have data on the survival probability of different Drosophila strains from egg to the third larval stage in one vector, $\vec{v}_1$, and from the third larval stage to pupation in a second vector, $\vec{v}_2$, then the total survival probability from egg through pupation could be represented as $\vec{v}_1 \circ \vec{v}_2$.

Another multiplication operator is the "*Kronecker product*" of two matrices, represented as $\mathbf{A} \otimes \mathbf{B}$. If **A** is an $r \times c$ matrix and **B** is an $s \times d$ matrix, then the Kronecker product is an $(r\,s) \times (c\,d)$ matrix, obtained by replacing the ij^{th} element of **A** with the matrix **B** times the scalar A_{ij}. As an example, consider two matrices that represent fitnesses of diploid individuals at genes A and B:

$$\mathbf{W_A} = \text{Fitnesses at gene } A = \begin{pmatrix} W_{AA} & W_{Aa} \\ W_{aA} & W_{aa} \end{pmatrix} \qquad \mathbf{W_B} = \text{Fitnesses at gene } B = \begin{pmatrix} W_{BB} & W_{Bb} \\ W_{bB} & W_{bb} \end{pmatrix}$$

where each row specifies the allele inherited from the mother (A or a; B or b) and each column specifies the allele inherited from the father. If the fitness of an individual is the product of the fitnesses predicted by each gene ("multiplicative selection"), we could represent the fitness of every genotype using the Kronecker product:

$$(\mathbf{A} \otimes \mathbf{B}) = \begin{pmatrix} W_{AA}\,W_{BB} & W_{AA}\,W_{Bb} & W_{Aa}\,W_{BB} & W_{Aa}\,W_{Bb} \\ W_{AA}\,W_{bB} & W_{AA}\,W_{bb} & W_{Aa}\,W_{bB} & W_{Aa}\,W_{bb} \\ W_{aA}\,W_{BB} & W_{aA}\,W_{Bb} & W_{aa}\,W_{BB} & W_{aa}\,W_{Bb} \\ W_{aA}\,W_{bB} & W_{aA}\,W_{bb} & W_{aa}\,W_{bB} & W_{aa}\,W_{bb} \end{pmatrix},$$

where now each row specifies the gamete type inherited from the mother (*AB*, *Ab*, *aB*, or *ab*) and each column specifies the gamete type inherited from the father. This example illustrates the book-keeping benefits of the Kronecker product; it would be easy to write down the fitness matrix for any arbitrary number of loci as $(\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C} \otimes \ldots \otimes \mathbf{N})$, but it would be tedious to write this matrix out by hand!

Advanced linear algebra texts provide further information about these alternative products and their properties (see Further Reading).

Supplementary Material P2.2: Calculating the determinant or the eigenvalues of a larger matrix

It can be cumbersome to find the determinant or the eigenvalues of a large matrix. Even *Mathematica* and other software packages can run into difficulties working with large matrices involving several parameters. Fortunately, there are operations that can be used to simplify the form of any matrix. These operations are known as the elementary row and column operations, and they have a known effect on the determinant. The elementary row operations are:

Operation 1: Two rows of a matrix can be interchanged

$\Rightarrow$The determinant switches sign

Operation 2: A row can be multiplied by any non-zero and finite constant, c

$\Rightarrow$The determinant is multiplied by c

Operation 3: A row times any constant can be added to or subtracted from another row

$\Rightarrow$The determinant is unchanged

The elementary column operations are similar, with "row" replaced by "column" in the above rules.

By repeatedly applying these operations, we can turn any matrix, **M**, into a triangular matrix, **M'**. The determinant of the triangular matrix is the product of the diagonal elements, and we can use the above rules to relate Det(**M'**) back to the determinant of the original matrix, Det(**M**). For example, let's turn the matrix $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ into a triangular matrix. The easiest way to do this is to use elementary row operation 3. Take the second row times b/d and subtract it from the first row to get $\mathbf{M'} = \begin{pmatrix} a - c\frac{b}{d} & 0 \\ c & d \end{pmatrix}$. The constant, b/d, was chosen to make the zero appear on the top right: the numerator was set to the term that we want to disappear (here, b) and the denominator was set to the term in the same column from the row to be subtracted (here, d).

Because elementary row operation 3 does not affect the determinant, Det(**M'**) equals Det(**M**). And, indeed, both determinants are $a\,d - b\,c$ (convince yourself of this).

These operations are also helpful in finding the eigenvalues of a matrix, **M**, which involves taking the determinant of $(\mathbf{M}-\lambda\mathbf{I})$ and setting it to zero. For example, consider a model of three equally-fit alleles, whose frequencies at a locus are p_1, p_2, and p_3. If each allele mutates at rate 2μ per generation and is converted equally to the other two alleles, then the changes in allele frequency per generation can be written in matrix form as:

$$\begin{pmatrix} p_1(t+1) \\ p_2(t+1) \\ p_3(t+1) \end{pmatrix} = \begin{pmatrix} 1-2\mu & \mu & \mu \\ \mu & 1-2\mu & \mu \\ \mu & \mu & 1-2\mu \end{pmatrix} \begin{pmatrix} p_1(t) \\ p_2(t) \\ p_3(t) \end{pmatrix}$$

The eigenvalues of this matrix can be used to predict the allele frequency trajectories (see Chapters 7 – 9). To find these eigenvalues, we must find the determinant of:

$$(\mathbf{M}-\lambda\mathbf{I}) = \begin{pmatrix} 1-2\mu-\lambda & \mu & \mu \\ \mu & 1-2\mu-\lambda & \mu \\ \mu & \mu & 1-2\mu-\lambda \end{pmatrix}$$

Calculating the determinant of a 3×3 and factoring is a bit tedious, so let's use the elementary row and column operations. It's always a good idea at this point to look for terms that are similar and that can be subtracted from each other without making the matrix horribly ugly. In this case, we'll start by subtracting the second column from the third column (elementary column operation 3):

$$(\mathbf{M}-\lambda\mathbf{I})' = \begin{pmatrix} 1-2\mu-\lambda & \mu & 0 \\ \mu & 1-2\mu-\lambda & -(1-3\mu-\lambda) \\ \mu & \mu & 1-3\mu-\lambda \end{pmatrix}$$

Next we'll add the third row to the second row (elementary row operation 3):

$$(\mathbf{M}-\lambda\mathbf{I})'' = \begin{pmatrix} 1-2\mu-\lambda & \mu & 0 \\ 2\mu & 1-\mu-\lambda & 0 \\ \mu & \mu & 1-3\mu-\lambda \end{pmatrix}$$

Finally, we'll multiply the first row by $1-\mu-\lambda$ (elementary row operation 2) and then subtract off the second row times μ to get:

$$(\mathbf{M}-\lambda\mathbf{I})''' = \begin{pmatrix} (1-\mu-\lambda)(1-2\mu-\lambda)-2\mu^2 & 0 & 0 \\ 2\mu & 1-\mu-\lambda & 0 \\ \mu & \mu & 1-3\mu-\lambda \end{pmatrix}$$

Because $(\mathbf{M}-\lambda\mathbf{I})'''$ is a lower triangular matrix, its determinant is the product of the diagonal elements, which after factoring equals $(1-\lambda)(1-3\mu-\lambda)^2(1-\mu-\lambda)$. In all of these calculations, we've only performed one operation that altered the determinant, which was when we multiplied a row by $1-\mu-\lambda$ (elementary row operation 2). This operation caused the determinant of $(\mathbf{M}-\lambda\mathbf{I})'''$ to be $(1-\mu-\lambda)$ times greater than the determinant of $(\mathbf{M}-\lambda\mathbf{I})$. Consequently, we can infer that the determinant of $(\mathbf{M}-\lambda\mathbf{I})$ is $(1-\lambda)(1-3\mu-\lambda)^2$. Setting this determinant to zero and solving for λ correctly identifies the three eigenvalues of the matrix, **M**, as 1, 1 – 3 μ, and 1 – 3 μ. Finding an eigenvalue that equals one informs us that there is some feature to the model that is invariant over time. In this model, we can identify this feature from the fact that the allele frequencies must always sum to one, $p_1 + p_2 + p_3 = 1$.

A word of caution: Avoid applying operation 3 using a constant whose denominator involves λ. This constant will not necessarily be finite for some choices of λ and can cause errors in the calculation of the eigenvalues.

From these examples, it might not be too obvious why we should bother with elementary row and column operations. After all, finding the determinant of a 2×2 matrix or even a 3×3 matrix isn't too cumbersome, even if the matrix isn't triangular. One advantage that we gained in this last example is that factoring was more straightforward than it would have been had we calculated the determinant using Rule P2.16. Where these operations really come in useful, however, is with higher dimensional matrices, where we can often "see" which elementary row operations can simplify a matrix in a way that a computer cannot.

Although we've discussed using elementary row and column operations to turn a matrix into a triangular matrix, you won't always have to go this far. If you can turn your original matrix into a block triangular matrix, then you can use Rule P2.21 to calculate the determinant from the diagonal sub-matrices. This helps enormously, for example, if the diagonal sub-matrices are two 2×2 matrices, whose determinants are much easier to calculate than the determinant of a 4×4 matrix! Alternatively, if you can use the elementary row and column operations to make all the non-diagonal elements in a row (or column) equal zero, then you can use Rule P2.22 to reduce the size of the matrix by one dimension, which is certainly a step in the right direction.

Exercise SP2.2.1: Use elementary row and column operations to find the determinant of the following matrix:

$$\mathbf{M} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 4 \\ 2 & 2 & 1 \end{pmatrix}$$ [Answer[i]]

Exercise SP2.2.2: A continuous-time version of the mutational model is:

$$\begin{pmatrix} dp_1/dt \\ dp_2/dt \\ dp_3/dt \end{pmatrix} = \begin{pmatrix} -2\mu & \mu & \mu \\ \mu & -2\mu & \mu \\ \mu & \mu & -2\mu \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

Use elementary row and column operations to find the eigenvalues of this rate matrix. How do these eigenvalues differ from the discrete-time model? [Answer[ii]]

Supplementary Material P2.3: Solving a system of linear equations using the method of row reduction

Many problems are related to solving a system of linear equations for a set of unknown variables. For example, we might wish to solve:

$$\mathbf{M}\,\vec{n} = \vec{b}$$

for the unknown vector $\vec{n}$. Here we describe a technique to solve this problem. First, create an "augmented matrix" by adding the vector $\vec{b}$ to the end of the matrix, **M**:

$$\left(\begin{array}{cccc|c} m_{11} & m_{12} & \dots & m_{1d} & b_1 \\ m_{21} & m_{22} & \dots & m_{2d} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{d1} & m_{d2} & \dots & m_{dd} & b_d \end{array}\right),$$

where we use the straight line through the matrix to help us remember that we've augmented **M** with $\vec{b}$. The augmented matrix is just a compact way of writing the system of equations where the first column gives the terms multiplying n_1, the second column gives the terms multiplying n_2, etc. Any solution vector, $\vec{n}$, that satisfies the equations represented by the augmented matrix will also work after applying an elementary row operation:

Operation 1: Two rows of the augmented matrix can be interchanged

Operation 2: A row can be multiplied by any non-zero constant, c

Operation 3: A row times any constant can be added to or subtracted from another row

A solution that causes the original equations to hold true would obviously satisfy the equations if we reordered them (operation 1). Furthermore, multiplying a true equation on both sides by a non-zero constant (operation 2) cannot make it false, nor will subtracting one true equation from another (operation 3).

The method of row reduction involves applying these elementary row operations to the augmented matrix until the augmented matrix is in the following form (known as the "reduced row-echelon form of the matrix"):

– The first non-zero element of any row is one (call this the leading term)

– The leading term in each row is to the right of the leading terms of the previous rows

– All other elements in a column in which a leading term is found are zeros

– If row reduction leads to a row of zeros, this row is moved to the bottom of the matrix

For example, if the matrix **M** is invertible, then our goal is to produce an augmented matrix:

$$\left(\begin{array}{cccc|c} 1 & 0 & \dots & 0 & f_1 \\ 0 & 1 & \dots & 0 & f_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & f_d \end{array}\right),$$

where f_i are functions of the b_i. This augmented matrix is equivalent to $\mathbf{I} \cdot \vec{n} = \vec{f}$, which indicates that $\vec{n} = \vec{f}$ is the solution that we sought.

To accomplish this procedure, follow these steps:

Step 1: Set $i = 1$ and $j = 1$.

Step 2: Make sure that the j^{th} element in row i is non-zero. If not, swap the first row with another row. If no row contains an element in the j^{th} column, then move over one column to the right (set j to $j + 1$).

Step 3: Divide the i^{th} row by the j^{th} entry in the row, making the leading term in the row equal one.

Step 4: From each of the remaining rows, subtract off the product of the j^{th} entry in each row multiplied by the i^{th} row that we have just calculated. This will create a series of zeros below the leading term of one in the i^{th} row and j^{th} column.

Step 5: Set i to $i + 1$ and j to $j + 1$ and return to Step 2. Repeat until the last row is reached.

Step 6: To create a series of zeros above the leading term in each row, determine the last row with a leading term of one. Set i to the number of this row and j to the column containing the leading term.

Step 7: From each of the rows 1 through i - 1, subtract off the product of the j^{th} entry multiplied by the i^{th} row.

Step 8: Set i to $i – 1$ and j to $j – 1$. If this term does not equal one, then set j to $j – 1$ again until the leading term is one. Return to Step 7. Repeat until the first row is reached.

This all sounds very abstract, but the procedure is pretty straightforward once you get used to it. As a simple numerical example, let's solve the following system of equations:

$$\begin{aligned} 2n_1 + n_2 + n_3 &= 2 \\ 2n_1 + 3n_2 + 5n_3 &= 4, \\ 2n_2 + 3n_3 &= 1 \end{aligned}$$

which can be written in matrix form as:

$$\begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 5 \\ 0 & 2 & 3 \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix},$$

and in augmented matrix form as:

$$\left(\begin{array}{ccc|c} 2 & 1 & 1 & 2 \\ 2 & 3 & 5 & 4 \\ 0 & 2 & 3 & 1 \end{array}\right)$$

Starting with $i = j = 1$, we find that there is a non-zero element in the first row of the first column, so we can move to Step 3 and divide the first row by the (i, j) element, 2:

$$\left(\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1 \\ 2 & 3 & 5 & 4 \\ 0 & 2 & 3 & 1 \end{array}\right)$$

Now turn to Step 4. From the second row, we subtract off the j^{th} element in the second row (= 2) multiplied by the first row to get:

$$\left(\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1 \\ 0 & 2 & 4 & 2 \\ 0 & 2 & 3 & 1 \end{array}\right)$$

The j^{th} (=1) element in the third row is already zero, so we can proceed to Step 5, setting $i = j = 2$ and going back to Step 2. Again, we find that there is a non-zero element in the second row and second column, so we can move to Step 3 and divide this second row by the element in this position (= 2):

$$\left(\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 2 & 3 & 1 \end{array}\right)$$

Now turn to Step 4. From the third row, we subtract off 2 (the j^{th} element in the third row, where now $j = 2$) times the second row:

$$\left(\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -1 & -1 \end{array}\right)$$

Finally, we can multiply the last row by –1, to give it a leading term of one as well:

$$\left(\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

Great. Now we can proceed to Step 6 and place zeros on top of the leading terms. The last row with a leading term in it is the third row, so we set $i = j = 3$. We then subtract off twice the third row from the second row to get a zero in the third column of the second row:

$$\left(\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right),$$

and subtract off 1/2 times the third row from the first row:

$$\left(\begin{array}{ccc|c} 1 & 1/2 & 0 & 1/2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

Step 8 is next. We move up to the leading term in the second row, setting $i = j = 2$. We need only subtract off 1/2 times the second row from the first row to finish the procedure:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

This is now in reduced row-echelon form and we can read off the answer from the column on the right-hand side of the straight line: $n_1 = 1$, $n_2 = -1$, and $n_3 = 1$. Plugging this solution into the original equation:

$$\begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 5 \\ 0 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}$$

confirms that we have indeed found the solution.

The method of row reduction can also be used to find the inverse of a matrix. Now, the equation that we wish to solve is:

$$\mathbf{M}\,\mathbf{M}^{-1} = \mathbf{I}$$

for the unknown matrix $\mathbf{M}^{-1}$. The "augmented matrix" now involves a matrix on the right-hand side of the straight line too:

$$\left(\begin{array}{cccc|cccc} m_{11} & m_{12} & \ldots & m_{1d} & 1 & 0 & \ldots & 0 \\ m_{21} & m_{22} & \ldots & m_{2d} & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{d1} & m_{d2} & \ldots & m_{dd} & 0 & 0 & \ldots & 1 \end{array}\right),$$

At this point, however, the method is exactly the same, just follow the steps outlined above. If matrix **M** is invertible (i.e., its determinant is not zero), then by row reduction you will eventually reduce the matrix on the left of the straight line to the identity matrix. At that point, the matrix to the right of the straight line is the inverse of **M**.

Exercise SP2.3.1: Use the method of row reduction to solve the following set of equations:

$$\begin{aligned} x + y + z &= 6 \\ x - y + z &= 2 \\ x - y - z &= 4 \end{aligned}$$

[Answer[iii]]

Exercise SP2.3.2: Use the method of row reduction to find the inverse of the following matrix:

$$\mathbf{M} = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

[Answer[iv]]

---

[i] ANSWER: Det(**M**) = 4.

[ii] ANSWER: The eigenvalues are 0, – 3 μ, and – 3 μ. Each of these eigenvalues is one less than the eigenvalues for the discrete-time model. This relationship is true whenever the continuous-time rate matrix is just the discrete-time transition matrix minus the identity matrix.

[iii] ANSWER: $x = 5$, $y = 2$, and $z = -1$.

[iv] ANSWER: $\mathbf{M}^{-1} = \begin{pmatrix} 1 & 1 & -3 \\ 0 & 1 & -1 \\ -1 & -2 & 5 \end{pmatrix}$.