Summary/Review of Matrix Algebra

Introduction to Matrices
  Descriptors & objects, Linear algebra, Order
Association Matrices
  R- and Q-mode
Special Matrices
  Trace, Diagonal, Identity, Scalars, Transpose
Vectors & Scaling
Matrix Addition and Multiplication
Determinants
Ranks
Inversion
Eigenvalues and Eigenvectors

Matrix Algebra

Matrix algebra is especially well suited to ecology and evolutionary biology (EEB) because most data sets are recorded in a matrix format (rows & columns).

Matrix notation provides an elegant and compact representation of ecological information and matrix algebra allows operations on whole data sets to be performed.

Multidimensional methods are virtually impossible to understand or conceptualize without resorting to matrix algebra.

Table Structure of Data

EEB data are generally recorded in a table (spreadsheet) where each column *j* corresponds to a descriptor y_j (species present in the sampling unit, physical trait, etc.) and each object x_i (sampling site, location, individual, etc.).

Thus, in each cell (*i*,*j*) of the table is found the state taken by object *i* for descriptor *j*:

Descriptors

Objects	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$	$\mathbf{y}_4$
$\mathbf{x}_1$	y_{11}	y_{12}	y_{13}	y_{14}
$\mathbf{x}_2$	y_{21}			
$\mathbf{x}_3$	y_{31}			
$\mathbf{x}_4$	y_{41}			

Objects, Descriptors, & Subscripts

Objects are usually denoted as a bold-face lowercase letter **x**, with a subscript *i* varying from 1 to *n*, referring to object x_i.

Similarly, descriptors will be noted by a bold-face lowercase letter **y** subscripted with a *j*, with *j* taking values from 1 to *p*, referring to descriptor y_j.

When considering two data sets simultaneously, we can use a second set of subscripts *k* from 1 to *m*.

Following this logic, each value in a data matrix can be denoted with a double subscript, the first subscript denoting the object being described and the second subscript the descriptor. Thus y_{83}, the the value taken by object 8 for descriptor 3.

Objects vs. Descriptors

It is not always obvious which are the objects and which are the descriptors.

For example, different sampling sites may be studied with respect to the different species found there. In contrast, when studying behavior or taxonomy, the organisms themselves are the subjects and a descriptor may be the type of habitat or site of occurrence.

One must be unambiguous and define what the objects and descriptors are *a priori*.

R-mode vs. Q-mode

The distinction between objects and descriptors is not just theoretical:

If you analyze the relationship among descriptors for the set of objects under study, you are conducting and R-mode analysis.

A Q-mode study examines the relationships among objects given a set of descriptors.

The mathematical techniques that are appropriate for one mode of analysis are NOT appropriate for the other mode.

Linear Algebra

A table of data as described above is an array of numbers referred to as a matrix. The branch of mathematics dealing with matrices is linear algebra.

Matrix **Y** (usually boldface, capital letter designation) is a rectangular ordered array of numbers y_{ij}, set in *n* rows and *p* columns:

$$\mathbf{Y} = [y_{ij}] = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{bmatrix}$$

Matrix Order

When the order (or dimension, or format) of the matrix must be specified, a matrix order ($n \times p$), which contains $n \times p$ elements, is written $\mathbf{Y}_{np}$. The elements of the matrix **Y** are denoted as $\mathbf{y}_{ij}$, with subscripts referring to rows and columns.

In linear algebra, ordinary numbers are referred to as scalars to distinguish them from matrices.

Matrix Form

Matrices can take many forms: rectangular, square, row or column.

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$$ Square Matrix

Row matrix (sub-matrix, row 1)

$$[y_{11} \quad y_{12} \quad \cdots \quad y_{1p}]$$

$$\begin{bmatrix} y_{12} \\ y_{22} \\ \vdots \\ y_{n2} \end{bmatrix}$$ Column matrix (sub-matrix, col 2)

Matrix Notation

Matrix notation simplifies the writing of data sets.

It also corresponds to the way computers works and the way programming languages interpret arrays of data.

Almost all computer packages permit (or encourage) data to be entered as matrices.

So, matrices are a VERY convenient approach to handling multivariate data!

Association Matrices

Two important matrices may be derived from an ecological data matrix:

The association matrix among objects
The association matrix among descriptors

Any association matrix is denoted as **A**, and contains the elements $\mathbf{a}_{ij}$.

We will spend a whole lecture on association matrices later in the course, but you only need to understand their *raison d'être* here.

Association

Using data from our previous matrix **Y**, one can examine relationships between the first two objects $\mathbf{x}_1$ and $\mathbf{x}_2$. In order to do so, the first and second rows of the matrix **Y**:

$$[y_{11} \quad y_{12} \quad \dots \quad y_{1p}] \textit{ and } [y_{21} \quad y_{22} \quad \dots \quad y_{2p}]$$

Are used to calculate a measure of association or similarity, to assess the degree of resemblance between the two objects.

This measure quantifies the strength of association between the two rows, and is denoted a_{12}.

Association Matrix

In the same way, all of the other pairwise associations between objects can be calculated.

The coefficients of association for all pairs of objects can be recorded in a table and ordered in such a way as to use for subsequent calculations. This is referred to as the association matrix **A** among objects:

$$\mathbf{A}_{nn} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Characteristics of the Association Matrix

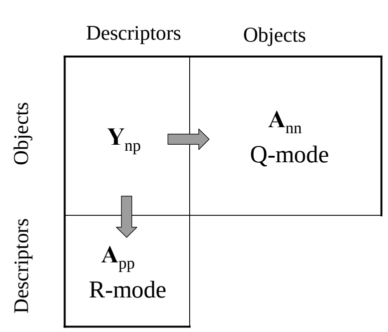
There are a couple of important characteristic of the association matrix:

(1) It is always square (same number of rows & columns)--this value being n, the number of objects. Thus, the total number of elements in the matrix is n^2.

(2) It is almost always symmetric, with elements in the upper hemi-matrix being identical to the elements in the lower hemi-matrix (i.e., $a_{ij} = a_{ji}$).

(3) The main diagonal is always unity (i.e., $a_{11} = a_{11}$).

Important Take-home Point



Special Matrices

As we just discussed, matrices with an equal number of rows and columns are called square matrices. **Y** matrices are rarely square, but the resulting **A** matrices are by definition.

As will be seen shortly, only square matrices can be used to compute a determinant, an inverse, and eigenvalues or eigenvectors.

As a corollary, all of these operations can always be carried out on **A** matrices.

Terminology for Square Matrices

For the square matrix **B** (of order n × n), the <u>diagonal elements</u> are those with identical subscripts (e.g., b_{11}, b_{22}, etc.) for rows and columns. They are located on the <u>main diagonal</u> (by convention from upper left to lower right).

$$\mathbf{B}_{nn} = [b_{ij}] = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix}$$

The sum of the diagonal elements is called the <u>trace</u> of the matrix.

Diagonal Matrices

A <u>diagonal matrix</u> is a form of square matrix where all of the non-diagonal elements are zero. For example:

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Diagonal matrices that contain values coming from a vector $[x_i]$ are noted as **D**(x). We will see special cases of this later such as the diagonal matrix of standard deviations **D**(σ) and the diagonal matrix of eigenvalues **D**(λ), often noted simply as **Λ** (cap. lambda) .

Diagonal Matrices

A <u>diagonal matrix</u> where all the diagonal elements are equal to unity is called a unit matrix or an identity matrix. It is denoted as **D**(1) or **I**:

$$\mathbf{D}(1) = \mathbf{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

This matrix is important because it plays the same role, in matrix algebra, as the number 1 in ordinary algebra; i.e., it is a neutral element (e.g., **IB** = **BI** = **B**).

Scalar Matrices

Similarly, a <u>scalar matrix</u> is a form of a diagonal matrix. A unit matrix multiplied by a scalar (e.g., 7) would contain 7's along the main diagonal:

$$7\mathbf{I} = \begin{bmatrix} 7 & 0 & \dots & 0 \\ 0 & 7 & & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 7 \end{bmatrix}$$

Null and Triangular Matrices

A matrix, square or rectangular, whose elements are all zero is called a <u>null matrix</u> or zero matrix and is denoted as **0** or [0].

A square matrix with all elements above (or below) the main diagonal being zero is called a lower (or upper) <u>triangular matrix</u>. These matrices are very important in matrix algebra because *their determinant is equal to the product of all terms on the main diagonal*.

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$

Transposed Matrices

The transpose of matrix **B** with format ($n \times p$) is denoted as **B'** and is a new matrix of format ($p \times n$) in which $b'_{ij} = b_{ji}$. In other words, the rows of one matrix become the columns of the other:

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} \quad \textit{and} \quad \mathbf{B'} = \begin{bmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{bmatrix}$$

Transposition is a very important operation in linear algebra, and also in ecology where **Y** is often transposed to **Y'** to study the relationships among descriptors after the relationships among objects have been analyzed.

Symmetric Matrices

A square matrix which is identical to its transpose is symmetric. This is the case when corresponding terms b_{ij} and b_{ji}, on either side of the diagonal, are equal. For example:

$$\begin{bmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{bmatrix}$$

Vectors

Another matrix of special interest is the column matrix, of format ($n \times 1$), which is also known as a vector. Row vectors are also possible of ($1 \times p$) format.

A column vector is noted as follows:

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

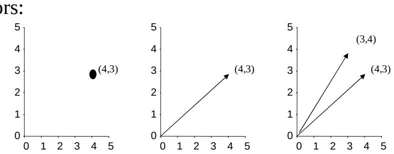
Vectors

A vector generally refers to a directed line segment, forming a mathematical entity on which operations can be performed.

More formally, a vector is defined as an ordered *n*-tuple of real numbers; i.e., a set of n numbers with a specified order. The *n* numbers are the co-ordinates of a point in *n*-dimensional Euclidean space, which may be seen as the end-point of a line segment starting at the origin.

Vectors

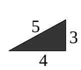
For example, the vector [4 3] is an ordered doublet (or 2-tuple) of two real numbers (4, 3), which may be represented in a two-dimensional Euclidean space. This same point may also be seen as the end-point of a line segment starting at the origin (0, 0). The vector [3 4] is different and stresses the ordered nature of vectors:



Vectors

Using the <u>Pythagorean Theorem</u>, it is now easy to calculate the <u>length</u> (or <u>norm</u>) of any vector. For example, the length of the previous vector [4 3] is the hypotenuse of a right triangle with a base of 4 and a height of 3. The length of the vector is thus:

$$\sqrt{4^2 + 3^2} = 5$$



This is also the same length of vector [3 4].

The norm of vector b is denoted as ‖b ‖.

Vectors

The comparison of different vectors, as to their directions, often requires an operation called <u>scaling</u>. In a scaled vector, all values are divided by the same characteristic value.

A special type of scaling is called <u>normalization</u>. In this type of scaling, every matrix element is divided by the length of the vector:

$$\begin{bmatrix} 4 \\ 3 \end{bmatrix} \rightarrow \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}$$

The importance of this is that the length of a normalized vector is equal to unity. Using the Pythagorean theorem, confirm for yourself that vector [4/5 3/5] = 1.

Generalization

The doublet example can be generalized to any *n*-tuple (b_1, b_2, …, b_n) which specifies *n*-dimensional space. The length of the vector is:

$$\sqrt{b_1^2 + b_2^2 + \ldots + b_n^2}$$

And the corresponding normalized vector is:

$$\begin{bmatrix} b_1 \,/\, \sqrt{b_1^2 + b_2^2 + \ldots + b_n^2} \\ b_2 \,/\, \sqrt{b_1^2 + b_2^2 + \ldots + b_n^2} \\ \vdots \\ b_n \,/\, \sqrt{b_1^2 + b_2^2 + \ldots + b_n^2} \end{bmatrix} = \frac{1}{\sqrt{b_1^2 + b_2^2 + \ldots + b_n^2}} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Matrix Addition

Matrices of the *same order* may be added = <u>matrix addition</u>.

Suppose you sample 5 sites and measure the abundance of 3 species of fish at 3 times during the summer and want to know the whole-summer total?

$$\begin{array}{l} \\ \text{Site-1} \\ \text{Site-2} \\ \text{Site-3} \\ \text{Site-4} \\ \text{Site-5} \end{array} \begin{array}{c} \text{June} \\ \begin{bmatrix} 1 & 5 & 35 \\ 14 & 2 & 0 \\ 0 & 31 & 67 \\ 96 & 110 & 78 \\ 0 & 0 & 0 \end{bmatrix} \end{array} + \begin{array}{c} \text{July} \\ \begin{bmatrix} 15 & 23 & 10 \\ 54 & 96 & 240 \\ 0 & 3 & 9 \\ 12 & 31 & 27 \\ 8 & 14 & 6 \end{bmatrix} \end{array} + \begin{array}{c} \text{August} \\ \begin{bmatrix} 48 & 78 & 170 \\ 2 & 0 & 0 \\ 0 & 11 & 14 \\ 25 & 13 & 12 \\ 131 & 96 & 43 \end{bmatrix} \end{array} = \begin{array}{c} \text{Summer} \\ \begin{bmatrix} 64 & 106 & 215 \\ 70 & 98 & 240 \\ 0 & 45 & 90 \\ 133 & 154 & 117 \\ 139 & 110 & 49 \end{bmatrix} \end{array}$$

sp-1 sp-2 sp-3

NB: Why was site-5 included in June when no fish?

Scalar Product

Suppose in the previous example that there was differential efficiency in with the sampling unit (say a net) to capture the fish species. Suppose you estimated the efficiencies for species 1, 2, & 3 as 100%, 50%, and 25%, respectively. You could create correction factors of 1, 2, & 4, respectively. To obtain the total fish abundance estimate for site-1 in July:

$$[1 \quad 5 \quad 35]\begin{bmatrix}1\\2\\4\end{bmatrix} = 151 \Leftrightarrow (1\times1)+(5\times2)+(35\times4)$$

151 is the scalar product of the two vectors.

Matrix Multiplication

The result of a scalar product is a number which is equal to the sum of the products of those corresponding order numbers. The scalar product is usually designated by a dot (or by no symbol at all). For example:

$$\mathbf{bc} = \mathbf{b} \bullet \mathbf{c} = [b_1 \quad b_2 \quad \ldots \quad b_p]\begin{bmatrix}c_1\\c_2\\\vdots\\c_p\end{bmatrix} = b_1c_1 + b_2c_2 + \ldots + b_pc_p = \text{ a scalar}$$

NB: Only vectors with the same number of elements can be multiplied!

Orthogonal Vectors

In analytical geometry, it can be shown that the scalar product of two vectors obeys an important relationship:

$$\mathbf{b} \bullet \mathbf{c} = (\text{length of } \mathbf{b}) \times (\text{length of } \mathbf{c}) \times \cos\theta$$

When the angle between two vectors is $\theta = 90$, then $\cos \theta = 0$ and the scalar product $\mathbf{b} \bullet \mathbf{c} = 0$. As a consequence, two vectors whose scalar product is zero are said to be orthogonal (i.e., at right angles). This is an important property (as you will see). A matrix whose (column) vectors are all at right angles is called orthogonal.

Returning to the fish example, we can now multiply each monthly matrix with a correction vector, in order to compare total monthly fish abundances. This produces a product of a vector by a matrix:

$$\begin{bmatrix} 1 & 5 & 35 \\ 14 & 2 & 0 \\ 0 & 31 & 67 \\ 96 & 110 & 78 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1(1) & + & 5(2) & + & 35(4) \\ 14(1) & + & 2(2) & + & 0(4) \\ 0(1) & + & 31(2) & + & 67(4) \\ 96(1) & + & 110(2) & + & 78(4) \\ 0(1) & + & 0(2) & + & 0(4) \end{bmatrix} = \begin{bmatrix} 151 \\ 18 \\ 330 \\ 628 \\ 0 \end{bmatrix}$$

June

Generalization

The product of a vector by a matrix involves calculating, for each row of a matrix B, a scalar product with vector c.

This is possible iff the number of elements in the vector is the same as the number of columns in the matrix! The result is a *vector*. Generalized:

$$\mathbf{B}_{pq} \bullet \mathbf{c}_q = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & b_{22} & \dots & b_{2q} \\ \vdots & & & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pq} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_q \end{bmatrix} = \begin{bmatrix} b_{11}c_1 + & b_{12}c_2 + & \dots & b_{1q}c_q \\ b_{21}c_1 + & b_{22}c_2 + & \dots & b_{2q}c_q \\ \vdots & & & \vdots \\ b_{p1}c_1 + & b_{p2}c_2 + & \dots & b_{pq}c_q \end{bmatrix}$$

Product of Two Matrices

The product of two matrices is a simple extension of the product of a vector by a matrix. Matrix **C**, to be multiplied by **B**, is simply considered a set of column vectors (e.g., c_1, c_2, etc.). For example:

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & 1 \\ 1 & 2 & 1 \\ -1 & 3 & 2 \end{bmatrix} \quad \textit{and} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & -1 \end{bmatrix} \qquad \mathbf{C} = [\mathbf{d}\ \mathbf{e}]$$

Multiply **B** by vectors **d** and **e** which make up **C**.

Product of Two Matrices **BC**

$$\mathbf{Bd} = \begin{bmatrix} 1(1)+ & 0(2)+ & 2(3) \\ 3(1)+ & 1(2)+ & 1(3) \\ 1(1)+ & 2(2)+ & 1(3) \\ -1(1)+ & 3(2)+ & 2(3) \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 8 \\ 11 \end{bmatrix}$$

and

$$\mathbf{Be} = \begin{bmatrix} 1(2)+ & 0(1)+ & 2(-1) \\ 3(2)+ & 1(1)+ & 1(-1) \\ 1(2)+ & 2(1)+ & 1(-1) \\ -1(2)+ & 3(1)+ & 2(-1) \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ 3 \\ -1 \end{bmatrix}$$

thus

$$\mathbf{BC} = \begin{bmatrix} 7 & 0 \\ 8 & 6 \\ 8 & 3 \\ 11 & -1 \end{bmatrix}$$

NB: Matrices to be multiplied MUST be conformable, in other words, the number of columns of the first matrix must be equal to the number of rows in the second matrix.

The result will be a matrix with the same no. of rows as **B** and no. of columns as **C.**

Determinants

It is often necessary to transform a matrix in to a new one, in such a way that the information in the original matrix is preserved, while new properties which are essential to subsequent calculations are acquired.

Such matrices, which are linearly derived from the original matrix, are often referred to under the names inverse matrix, canonical form, etc.

The new matrix must have a minimum number of characteristics in common with the matrix from which it was derived...

The connection between the two matrices is a matrix function $f(\mathbf{B})$, whose properties are the following:

(1) The function must be multi-linear, which means that it should respond linearly to any change taking place in the rows or columns.

(2) Since the order of the rows and columns of the matrix is specified, the function should be able to detect, through alteration of signs, any change in position of rows or columns. If two rows or columns are identical, $f(\mathbf{B}) = 0$.

(3) $f(\mathbf{B})$ has a scalar associated with it; it is called the norm or value. The norm is calibrated in such a way that the value associated with the unit matrix **I** is 1; i.e., $f(\mathbf{I}) = 1$.

The determinant, is the only function which has these 3 properties, and it only exists for square matrices.

The determinant of matrix **B** is denoted as either det**B**, or more often, |**B**|:

$$|B| \equiv \begin{vmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & & & \vdots \\ b_{p1} & b_{p2} & \dots & b_{nn} \end{vmatrix}$$

Vertical lines are used to symbolize an "array" as opposed to a matrix *sensu stricto*.

The value of a determinant is calculated as the sum of all possible products containing one, and only one, element from each row and each column; these products receive a sign according to a well-defined rule:

$$|B| = \sum \pm (b_{1j_1}\ b_{2j_2}\ \dots b_{nj_n})$$

Where j_1, j_2, ... j_n, go through the n! permutations of the numbers 1, 2, ..., n. The sign depends upon the number of inversions, in the permutation considered, relative to the regular sequence 1, 2, ..., n: if the number of inversions is even, the sign is (+), if odd then the sign is (-).

Determinant of a Matrix of Order 2

For example, the determinant of a matrix of order 2 is calculated:

$$|B| = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = b_{11}b_{22} - b_{12}b_{21}$$

In accordance with the formal definition, the scalar so obtained is composed of 2!=2 products, each containing one, and only one, element from each row and each column.

The determinant of a matrix of order higher than 2 may be calculated using a variety of approaches. The two most common approaches are (1) expansion by minors and (2) pivotal condensation.

The latter is very calculation intensive and usually implemented on a computer. We will take a closer look at expansion by minors.

When looking for a determinant of order 3, a determinant of order 3-1 = 2 may be obtained by crossing out one row (*i*) and one column (*j*) of data. This lower order determinant is the minor associated with b_{ij}.

Crossing out row 1 and col 2

$$\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} \rightarrow \begin{vmatrix} b_{21} & b_{23} \\ b_{31} & b_{33} \end{vmatrix}$$

minor of b_{12}

When multiplied by $(-1)^{i+j}$, the minor becomes a cofactor. Thus, the cofactor of b_{12} is:

$$\text{cof } b_{12} = (-1)^{1+2} \begin{vmatrix} b_{21} & b_{23} \\ b_{31} & b_{33} \end{vmatrix} = -\begin{vmatrix} b_{21} & b_{23} \\ b_{31} & b_{33} \end{vmatrix}$$

Thus, going back to the determinant equation used earlier, expansion by elements of the first row gives:

$$|\mathrm{B}| = b_{11} \text{ cof } b_{11} + b_{12} \text{ cof } b_{12} + b_{13} \text{ cof } b_{13}$$

$$|\mathrm{B}| = b_{11}(-1)^{1+1} \begin{vmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{vmatrix} + b_{12}(-1)^{1+2} \begin{vmatrix} b_{21} & b_{23} \\ b_{31} & b_{33} \end{vmatrix} + b_{13}(-1)^{1+3} \begin{vmatrix} b_{21} & b_{22} \\ b_{31} & b_{32} \end{vmatrix}$$

Let's look at a numerical example:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix}$$

Order 3 Determinant Calculation Example

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix} = 1(-1)^{1+1}\begin{vmatrix} 5 & 6 \\ 8 & 10 \end{vmatrix} + 2(-1)^{1+2}\begin{vmatrix} 4 & 6 \\ 7 & 10 \end{vmatrix} + 3(-1)^{1+3}\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix}$$

or

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix} = 1(5\times10-6\times8)-2(4\times10-6\times7)+3(4\times8-5\times7) = -3$$

As you can see, the amount of calculations required to expand a determinant increases VERY quickly with increasing order n.

Systat 8.0: MATRIX DET procedure

Linear Dependency

A square matrix contains *n* vectors (rows or columns), which may be linearly independent or not. Two vectors are linearly dependent when the elements of one are proportional to those of the other. For example:

$$\begin{bmatrix} -4 \\ -6 \\ -8 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \text{ are linearly dependent since } \begin{bmatrix} -4 \\ -6 \\ -8 \end{bmatrix} = -2\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

Rank of a Square Matrix

The rank of a square matrix is defined as the number of linearly independent row (or column) vectors in the matrix. For example:

$$\begin{bmatrix} -1 & -1 & 1 \\ 3 & 0 & -2 \\ 4 & 1 & -3 \end{bmatrix}$$

(-2 × col 1) = col 2 + (3 × col 3)
or row 1 = row 2 - row 3
rank = 2

$$\begin{bmatrix} -2 & 1 & 4 \\ -2 & 1 & 4 \\ -2 & 1 & 4 \end{bmatrix}$$

(-2 × col 1) = (4 × col 2) = col 3
or row 1 = row 2 = row 3
rank = 1

Matrix Inversion

In algebra, division is expressed as either:

$c \div b$, c/b, $c(1/b)$ or $c\, b^{-1}$

In the last two expressions, division is replaced by multiplication with a reciprocal or inverse quantity.

In matrix algebra, division of **C** by **B** does not exist. The equivalent operation is multiplication of **C** with the inverse or reciprocal of matrix **B**.

The inverse of matrix **B** is denoted $\mathbf{B}^{-1}$ and the operation through which it is calculated is called inversion of matrix **B**.

Matrix Inversion

To serve its purpose, matrix $\mathbf{B}^{-1}$ must be unique and the relation $\mathbf{B}\mathbf{B}^{-1} = \mathbf{B}^{-1}\mathbf{B} = I$ must be satisfied. It can be shown that only square matrices have unique inverses.

To calculate the inverse of a square matrix **B**, the adjugate or adjoint matrix of **B** must first be defined. In the matrix of cofactors of **B**, each element b_{ij} is replaced by it cofactor. The adjugate matrix of **B** is the transpose of the matrix of the cofactors. Thus...

Matrix Inversion

$$\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix} \rightarrow \begin{bmatrix} \text{cof } b_{11} & \text{cof } b_{21} & \dots & \text{cof } b_{n1} \\ \text{cof } b_{12} & \text{cof } b_{22} & \dots & \text{cof } b_{n2} \\ \vdots & & & \vdots \\ \text{cof } b_{1n} & \text{cof } b_{2n} & \dots & \text{cof } b_{nn} \end{bmatrix}$$

matrix **B** adjugate matrix of **B**

Matrix Inversion

The inverse of matrix **B** is the adjugate matrix of **B** divided by the determinant |**B**|. The product of the matrix with its inverse should yield the unit matrix **I**.

$$\underbrace{\frac{1}{|\mathbf{B}|}\begin{bmatrix} \text{cof } b_{11} & \text{cof } b_{21} & \dots & \text{cof } b_{n1} \\ \text{cof } b_{12} & \text{cof } b_{22} & \dots & \text{cof } b_{n2} \\ \vdots & & & \vdots \\ \text{cof } b_{1n} & \text{cof } b_{2n} & \dots & \text{cof } b_{nn} \end{bmatrix}}_{\mathbf{B}^{-1}} \underbrace{\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix}}_{\mathbf{B}} = \mathbf{I}$$

All diagonal terms resulting from the multiplication of $\mathbf{B}\mathbf{B}^{-1}$ or $\mathbf{B}^{-1}\mathbf{B}$ are of the form $\Sigma\, b_{ij} \text{ cof } b_{ij}$ which is the expansion by minors of a determinant.

Each diagonal element consequently has the value of the determinant |**B**|. All other elements of the $\mathbf{B}^{-1}\mathbf{B}$ are sums of the products of the elements of a row with the corresponding cofactors of a different row. Each non-diagonal element is therefore null.

Therefore, it must follow that:

$$\mathbf{B}^{-1}\mathbf{B} = \frac{1}{|\mathbf{B}|}\begin{bmatrix} |\mathbf{B}| & 0 & \dots & 0 \\ 0 & |\mathbf{B}| & \dots & 0 \\ \vdots & & |\mathbf{B}| & \vdots \\ 0 & 0 & \dots & |\mathbf{B}| \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & 1 & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \mathbf{I}$$

An important point is that $\mathbf{B}^{-1}$ exists only if $|\mathbf{B}| \neq 0$. A square matrix with a null determinant is known as a singular matrix and it has no inverse. Matrices which can be inverted are known as nonsingular.

Let's return to the numerical example we used previously (since we already calculated it's determinant = -3):

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix} \quad \begin{bmatrix} 2 & 2 & -3 \\ 4 & -11 & 6 \\ -3 & 6 & 3 \end{bmatrix} \quad \begin{bmatrix} 2 & 4 & -3 \\ 2 & -11 & 6 \\ -3 & 6 & -3 \end{bmatrix} \quad -\frac{1}{3}\begin{bmatrix} 2 & 4 & -3 \\ 2 & -11 & 6 \\ -3 & 6 & -3 \end{bmatrix}$$

matrix   matrix of cofactors   transpose matrix   inverse of matrix

Systat 8.0: MATRIX INV procedure

Eigenvalues and Eigenvectors

There are many problems where the determinant and the inverse of the matrix can be used to provide simple and elegant solutions. Earlier in the course, you developed matrix operations up to this point to determine coefficients in multiple linear regression.

We will diverge to an alternative use, and a very important one in multivariate data analysis, and that is the derivation of an orthogonal form (i.e., a matrix whose vectors are at right angles) for a non-orthogonal symmetric matrix. This will provide the basis for most ordination methods, which we will examine next.

In EEB, data sets generally include a large number of variables, which are associated with one another (e.g., linearly correlated). The basic idea underlying many multivariate methods is to reduce this large number of intercorrelated variables to a smaller number of composite, but linearly independent variables, each explaining a different fraction of the observed variation.

First we need to deal with the mathematics behind the computation of eigenvalues and eigenvectors. In a subsequent lecture we will work with multidimensional EEB data.

Eigenvalues and Eigenvectors

Mathematically, the problem may be formulated as follows. Given a square matrix **A**, one wishes to find a diagonal matrix which is equivalent to **A**. In EEB, square matrices are most often symmetric association matrices, hence the use of the symbol **A**.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Eigenvalues and Eigenvectors

In matrix A, the terms located above and below the diagonal characterize the degree of association of either the objects, or the variables, with one another.

In the new matrix $\mathbf{\Lambda}$ (capital lambda) being sought, all the elements outside the diagonal are null:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_{11} & 0 & \dots & 0 \\ 0 & \lambda_{22} & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & \lambda_{nn} \end{bmatrix} = \begin{bmatrix} \lambda_{1} & 0 & \dots & 0 \\ 0 & \lambda_{2} & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & \lambda_{n} \end{bmatrix}$$

Eigenvalues and Eigenvectors

This new matrix Λ is called the matrix of eigenvalues*.

The new variables (eigenvectors) whose association are described by this matrix are thus linearly independent of one another.

Matrix Λ is known as the canonical form of matrix A; we will see the exact meaning of canonical in a future lecture.

*SYNONYMS: eigenvalue, characteristic root, latent root, eigenvector, characteristic vector, latent vector.

Computation

The eigenvalues and eigenvectors of matrix A are found from the equation:

$$\mathbf{Au}_i = \lambda_i \mathbf{u}_i$$

which allows one to compute the different values of λ_i and their associated eigenvectors $\mathbf{u}_i$.

If scalars λ_i and their associated vectors exist, then this eq. can be transformed in to:

$$\mathbf{Au}_i - \lambda_i \mathbf{u}_i = 0\text{; and factorizing } \mathbf{u}_i\text{: } (\mathbf{A} - \lambda_i)u_i = 0$$

Computation

For a matrix **A** of order *n*, the characteristic equation is a polynomial of degree *n*, whose solutions are the values λ_i.

When these values are found, it is easy to calculate the eigenvector $\mathbf{u}_i$ corresponding to each λ_i using our equation

$$(\mathbf{A} - \lambda_i)u_i = 0$$

If $\mathbf{u}_i$ has to be a null column vector, then we can introduce a unit matrix inside the brackets: $|\mathbf{A} - \lambda_i\mathbf{I}| = 0$

That is, the determinant of the difference between matrices **A** and $\lambda_i\mathbf{I}$ must be equal to zero for each λ_i. Resolving this equation yields the characteristic equation.

Numerical Example

Given the square symmetric matrix **A** : $\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$

The characteristic root is calculated as:

$$\left| \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0$$

$\therefore$

$$\left| \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$$

and thus,

$$\begin{vmatrix} 2\text{-}\lambda & 2 \\ 2 & 5\text{-}\lambda \end{vmatrix} = 0$$

The char. polynomial is found by expanding the determinant:

$$(2-\lambda)(5-\lambda) - 4 = 0$$

which gives

$$\lambda^2 - 7\lambda + 6 = 0$$

the characteristic polynomial.

Now, solve the characteristic polynomial for the two values. Recall from elementary algebra the solution for a quadratic equation using the quadratic formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

for $\lambda^2 - 7\lambda + 6 = 0$

the coefficients are :

$a = 1$

$b = 7$

$c = 6$

solving :

$$\lambda = \frac{-(-7) \pm \sqrt{7^2 - 4(1)(6)}}{2(1)}$$

$$= \frac{7 \pm 5}{2} = 6,1$$

$$\therefore \lambda_1 = 6, \lambda_2 = 1$$

The ordering of the two eigenvalues is arbitrary.

Systat 8.0: MATRIX EIGVAL procedure

Using the Systat matrix procedure, the eigenvalues can be computed in a similar manner and you get the same result of 6 & 1.

We now need to calculate the eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ corresponding to eigenvalues λ_1 and λ_2.

for $\lambda_1 = 6$

$$\left(\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} - 6\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} u_{11} \\ u_{22} \end{bmatrix} = 0$$

$$\begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} u_{11} \\ u_{22} \end{bmatrix} = 0$$

which is equivalent to the linear equations:

$-4u_{11} + 2u_{21} = 0$ and $2u_{11} - 1u_{21} = 0$

for $\lambda_2 = 1$

$$\left(\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} - 1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} u_{11} \\ u_{22} \end{bmatrix} = 0$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} u_{11} \\ u_{22} \end{bmatrix} = 0$$

which is equivalent to the linear equations:

$1u_{12} + 2u_{22} = 0$ and $2u_{12} - 4u_{22} = 0$

These sets of linear equations are always <u>indeterminate</u>. The solution is given by any point (vector) in the direction of the eigenvector being sought. To remove the indetermination, an arbitrary value is assigned to one of the elements of **u**, which specifies a particular vector.

For example, we might continue by assigning **u** and arbitrary value of 1 in each set.

given that	$\mathbf{u}_{11} = 1$	given that	$\mathbf{u}_{12} = 1$
it follows that	$-4\mathbf{u}_{11} + 2\mathbf{u}_{21} = 0$	it follows that	$1\mathbf{u}_{12} + 2\mathbf{u}_{22} = 0$
become	$-4 + 2\mathbf{u}_{21} = 0$	become	$1 + 2\mathbf{u}_{22} = 0$
so that	$\mathbf{u}_{21} = 2$	so that	$\mathbf{u}_{22} = -1/2$
Eigenvector $\mathbf{u}_1$ is $\therefore$	$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	Eigenvector $\mathbf{u}_2$ is $\therefore$	$\begin{bmatrix} 1 \\ -1/2 \end{bmatrix}$

NB: Values other than 1 could have been used; the resulting eigenvectors would differ only by the multiplication of a scalar.

This scalar difference WRT the eigenvectors is why they are generally <u>standardized</u>.

(1) One method is to assign the largest element of each vector a value of 1, and adjust the other elements accordingly.

(2) Another standardization method commonly used in principal components analysis is to make the length of each eigenvector ui equal to the square root of its eigenvalue ($\sqrt{\lambda_i}$).

(3) The most common procedure is to make the lengths of all the eigenvectors equal to 1 (i.e., **u'u** = 1). This is achieved by dividing each element of a vector by the length of this vector, i.e., the square root of the sum of the squares of the elements in the vector.

Continuing with our numerical example, the two eigenvectors can thus be normalized:

$$\begin{bmatrix}1\\2\end{bmatrix} \downarrow \begin{bmatrix}1/\sqrt{5}\\2/\sqrt{5}\end{bmatrix} \qquad \begin{bmatrix}2\\-1\end{bmatrix} \downarrow \begin{bmatrix}2/\sqrt{5}\\-1/\sqrt{5}\end{bmatrix}$$

Since matrix A is symmetric, its eigenvectors must be orthogonal. This is easily verified as their product must equal zero, which is the condition for the two vectors to be orthogonal:

$$\mathbf{u}_1^{'}\mathbf{u}_2 = \begin{bmatrix}1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}\begin{bmatrix}2/\sqrt{5}\\-1/\sqrt{5}\end{bmatrix} = 2/5 - 2/5 = 0$$

Systat 8.0: MATRIX CALL procedure

CALL EIGEN (B, C, A)

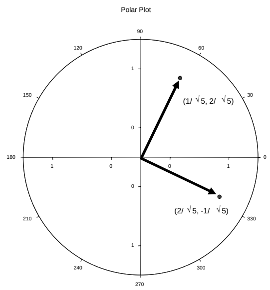
Takes data matrix A, calculates eigenvalues and eigenvectors and places in matrices B & C, respectively.

NB: 0.447 = 1/√5
0.894 = 2/√5
Same as worked example

Cartesian Space Solution

The normalized eigenvectors can now be plotted in the original system of coordinates, i.e., the Cartesian plane whose axes are the two original descriptors; the association between these descriptors is given by the matrix **A**.

This plot shows that indeed the angle between the eigenvectors is 90º (cos 90º = 0) and their lengths = 1.



Resolving the system of linear equations used to compute eigenvectors is greatly facilitated by matrix inversion. Defining matrix $\mathbf{C}_{nn} = (\mathbf{A} - \lambda_n \mathbf{I})$ allows the equation we have been using to be written in a more simplified form:

$$\mathbf{C}_{nn}\mathbf{u}_n = 0$$

Matrix $\mathbf{C}_{nn}$ contains all of the coefficients by which a given eigenvector $\mathbf{u}_n$ is multiplied. The system of equations is indeterminate, which prevents the inversion of **C** and calculation of **u**.

To remove the indetermination, it is sufficient to determine any one element of vector **u**. For example, one may arbitrarily decide that $u_1 = \alpha$ ($\alpha \neq 0$), then...

$$\begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & & & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix} \begin{bmatrix} \alpha \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

which can now be written as:

$$\begin{bmatrix} c_{11}\alpha + & c_{12}u_2 + & \dots & + c_{1n}u_n \\ c_{21}\alpha + & c_{22}u_2 + & \dots & + c_{2n}u_n \\ \vdots & & & \vdots \\ c_{n1}\alpha + & c_{n2}u_2 + & \dots & + c_{nn}u_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

so that

$$\begin{bmatrix} c_{12}u_2 & + & \dots & c_{1n}u_n \\ c_{22}u_2 & + & \dots & c_{2n}u_n \\ \vdots & & & \vdots \\ c_{n2}u_2 & + & \dots & c_{nn}u_n \end{bmatrix} = -\alpha \begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{bmatrix}$$

After setting $u_1 = \alpha$, the first column of matrix **C** is transferred to the right. The last n-1 rows are enough to defined a completely determined system.

The first row is then removed (not shown) to get a square matrix of order n-1, which can then be inverted.

The End

From here, we will use matrix methods to reduce the dimensionality of biological data ... stay tuned for PCA.