Bio Stats II Lecture 11: Matrix Algebra Review

Gavin Fay

02/28/2023

This Week...

2/28: Matrix Algebra review
3/01: Lab 7: Functions
3/02: Principal Component Analysis

Objectives

- ▶ Why Matrix Algebra?
- ▶ Review matrix operations
- ▶ Solving systems of linear equations using matrices
- ▶ Eigenanalysis

Why Matrix Algebra?

Most biological datasets are recorded in a matrix format (rows & columns).

Matrix notation gives elegant & compact representation of information.

Matrix algebra allows operations on whole data sets to be performed.

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

Branch of mathematics dealing with matrices is *linear algebra*.

Where are we going?

Linear regression, explain variability in Y as function of X.

Matrices

Matrix **Order** is the dimension of the matrix. e.g. a matrix order ($m \times n$) contains mn elements, has m rows and n columns:

$$A_{m,n} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

Matrix form:

Square matrix:

$$\begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \end{bmatrix}$$

Row matrix:

$$\begin{bmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,n} \end{bmatrix}$$

Column matrix:

$$\begin{bmatrix} y_{1,1} \\ y_{2,1} \\ \cdots \\ y_{m,1} \end{bmatrix}$$

Matrix Addition and Subtraction

If matrices are of the same order, they can be added. $A + B = C$

Addition is done element by element.

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} + \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} = \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} \end{pmatrix} = \begin{pmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{pmatrix}$$

e.g.

$$\begin{pmatrix} 1 & 4 \\ 5 & -3 \end{pmatrix} + \begin{pmatrix} -3 & 2 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 1-3 & 4+2 \\ 5+2 & -3+2 \end{pmatrix} = \begin{pmatrix} -2 & 6 \\ 7 & -1 \end{pmatrix}$$

Matrix addition is commutative: $A + B = B + A$

Matrix addition is associative: $(A + B) + C = A + (B + C)$

Matrix transposition

The *transpose* of a matrix A is denoted A^T or A'. Order subscripts are exchanged.

$$Q = \begin{bmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \\ q_{3,1} & q_{3,2} \end{bmatrix} \quad Q' = \begin{bmatrix} q_{1,1} & q_{2,1} & q_{3,1} \\ q_{1,2} & q_{2,2} & q_{3,2} \end{bmatrix}$$

$$w = \begin{bmatrix} 3 & -2 & 5 & 7 \end{bmatrix} \quad w' = \begin{bmatrix} 3 \\ -2 \\ 5 \\ 7 \end{bmatrix}$$

A matrix is *symmetric* if $A = A'$ (all symmetric matrices are square)

Transpose of a transpose is the original matrix: $(A')' = A$

For a scalar k, $(kA)' = kA'$

Multiplication

scalar multiplication:

$$A = \alpha \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} = \begin{bmatrix} \alpha a_{1,1} & \alpha a_{1,2} \\ \alpha a_{2,1} & \alpha a_{2,2} \end{bmatrix}$$

So for:

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

Matrix Multiplication

Consider A of order (m, n) and B of order (n, r).

Product $AB = C$ is order (m, r) with elements $c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$

To be able to do matrix multiplication, number of columns in first matrix must equal the number of rows in the second matrix.

Let

$$A = \begin{bmatrix} 4 & 2 & 6 \\ 2 & -1 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 2 & -2 \\ 1 & -5 \\ 3 & 2 \end{bmatrix}$$

$$C = AB = \begin{bmatrix} 4 \times 2 + 2 \times 1 + 6 \times 3 & 4 \times (-2) + 2 \times (-5) + 6 \times 2 \\ 2 \times 2 + (-1) \times 1 + 4 \times 3 & 2 \times (-2) + (-1) \times (-5) + 4 \times 2 \end{bmatrix}$$

$$C = AB = \begin{bmatrix} 4 \times 2 + 2 \times 1 + 6 \times 3 & 4 \times (-2) + 2 \times (-5) + 6 \times 2 \\ 2 \times 2 + (-1) \times 1 + 4 \times 3 & 2 \times (-2) + (-1) \times (-5) + 4 \times 2 \end{bmatrix}$$

$$= \begin{bmatrix} 8 + 2 + 18 & -8 - 10 + 12 \\ 4 - 1 + 12 & -4 + 5 + 8 \end{bmatrix} = \begin{bmatrix} 28 & -6 \\ 15 & 9 \end{bmatrix}$$

Matrix multiplication is *not* commutative.

BA is of order $(3,3)$. $AB \neq BA$

Matrix Multiplication

Other things to know:

Matrix multiplication is associative: $(AB)C = A(BC)$

Matrix multiplication is distributive: $A(B + C) = AB + AC$

Scalar multiplication commutative, associative, & distributive.

Transpose of a product: $(AB)' = B'A'$

Use exponentiation operator to denote repeated multiplication:
$A^3 = A \cdot A \cdot A$

Matrix multiplication

Sum-of-squares and cross-products

$$\mathbf{Y} - \mathbf{U} = \mathbf{D}$$

$$\begin{bmatrix} 1 & 4 & 10 \\ 2 & 6 & 15 \\ 3 & 8 & 20 \end{bmatrix} - \begin{bmatrix} 2 & 6 & 15 \\ 2 & 6 & 15 \\ 2 & 6 & 15 \end{bmatrix} = \begin{bmatrix} -1 & -2 & -5 \\ 0 & 0 & 0 \\ 1 & 2 & 5 \end{bmatrix}$$

$$\mathbf{D}^{\mathbf{T}}\mathbf{D} = \mathbf{S}$$

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -5 & 0 & 5 \end{bmatrix}\begin{bmatrix} -1 & -2 & -5 \\ 0 & 0 & 0 \\ 1 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 10 \\ 4 & 8 & 20 \\ 10 & 20 & 50 \end{bmatrix}$$

Population projection using Matrices

Leslie Matrix of population age structure:

Vector N_t is the numbers at age. i.e.

$$N_t = \begin{bmatrix} n_{0,t} \\ n_{1,t} \\ \cdots \\ n_{J,t} \end{bmatrix}$$

Life table matrix consists of the survival from one age to the next on the off-diagonals, and the fecundity of mature individuals on the first row.

$$X = \begin{bmatrix} 0 & f_1 & f_2 & \cdots & f_{J-1} & f_J \\ S_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & S_1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & S_2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & S_{J-1} & S_J \end{bmatrix}$$

Population projection

$$N_{t+1} = XN_t = \begin{bmatrix} n_{1,t} \cdot f_1 + n_{2,t} \cdot f_2 + \cdots + n_{J,t} \cdot f_J \\ n_{0,t} \cdot S_0 \\ n_{1,t} \cdot S_1 \\ \cdots \\ n_{J-1,t} \cdot S_{J-1} + n_{J,t} \cdot S_J \end{bmatrix}$$

Matrix Inverse

In scalar algebra, the inverse of a number is that number which, when multiplied by the original number, gives a product of 1. i.e. inverse of x is $1/x$, or x^{-1}

In matrix algebra, the inverse of a matrix is that matrix which, when multiplied by the original matrix, gives an identity matrix.

$$AA^{-1} = A^{-1}A = I$$

A matrix must be square to have an inverse, but not all square matrixes have an inverse.

(The identity matrix I is a matrix with 1's in the diagonal, and 0's in the off-diagonals)

There are several approaches to inverting matrices. We need to know what an inverse is in multivariate statistics, but it is not necessary to know how to compute it.

Determinant of a Matrix

For covariance and correlation matrices, the determinant of a matrix ($|A|$ or $det(A)$) is a scalar sometimes used to express the generalized variance of the matrix.

e.g. for a 2 x 2 matrix:

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Covariance matrices with small determinants indicate variables are highly correlated. Matrices with large determinants indicate variables that are independent.

Trace

The trace of a matrix, $tr(A)$, is used for square matrices and is the sum of the diagonal.

$$tr \begin{pmatrix} 2 & -1 & 3 \\ 1 & 4 & 2 \\ 8 & 0 & -2 \end{pmatrix} = 2 + 4 - 2 = 4$$

Useful matrices

Orthogonal matrices - must be square matrices

$$AA^T = I$$

i.e. inverse of an orthogonal matrix is its transpose.

Diagonal matrix - off-diagonals are 0.

Identity matrix - a diagonal matrix where the elements on the diagonal all $= 1$.

Regression using matrices

For regression through the origin ($\hat{y} = bx$), the regression coefficient (b) is the ratio of sum-of-products and sum-of-squares for x.

$$b = \frac{\sum_n x_i y_i}{\sum_n x_i^2}$$

More easily calculated as matrix operations. X & Y are data vectors, b is scalar.

$$b = (X^T X)^{-1} X^T Y \quad , \quad \hat{Y} = bX$$

$$\begin{bmatrix} \hat{y}_i \\ \cdots \\ \hat{y}_n \end{bmatrix} = b \begin{bmatrix} x_i \\ \cdots \\ x_n \end{bmatrix}$$

Multiple regression with matrices

expand X from univariate vector to multivariate matrix expand b from scalar to vector B of regression coefficients Intercepts are usually as 1st element of B, and 1st column of X is all 1's.

$$B = (X^T X)^{-1} X^T Y \quad , \quad \hat{Y} = XB$$

$$\begin{bmatrix} \hat{y}_i \\ \cdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} x_{i1} & \cdots & x_{ip} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} b_1 \\ \cdots \\ b_p \end{bmatrix}$$

Variance-Covariance Matrix

Variance is the average sum-of-squares of a variable (corrected for low sample size, n-1). Covariance is average cross-product of 2 variables. Covariance matrix is symmetric: - variance of each variable on the diagonal - paired covariances on the off diagonals

$$s_{y_1}^2 = \frac{\sum_n (y_{1i} - \bar{y}_1)^2}{(n-1)}$$

$$s_{y_1 y_2} = \frac{\sum_n (y_{1i} - \bar{y}_1)(y_{2i} - \bar{y}_2)}{(n-1)}$$

$$C = \begin{bmatrix} s_{y_1}^2 & \cdots & s_{y_{1p}} \\ \cdots & \cdots & \cdots \\ s_{y_{p1}} & \cdots & s_{y_p}^2 \end{bmatrix}$$

Variance-Covariance matrix

$$\mathbf{D}^T\mathbf{D} = \mathbf{S}$$

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -5 & 0 & 5 \end{bmatrix} \begin{bmatrix} -1 & -2 & -5 \\ 0 & 0 & 0 \\ 1 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 10 \\ 4 & 8 & 20 \\ 10 & 20 & 50 \end{bmatrix}$$

$$\left(\frac{1}{n-1} \right)\mathbf{S} = \mathbf{C}$$

$$0.5 \begin{bmatrix} 2 & 4 & 10 \\ 4 & 8 & 20 \\ 10 & 20 & 50 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \\ 5 & 10 & 25 \end{bmatrix}$$

Correlation

Correlation is covariance of 2 variables divided by pooled variance of both.

$$r_{y_1 y_2} = \frac{s_{y_1 y_2}}{s_{y_1} s_{y_2}}$$

Eigenanalysis

Similar to the way an inverse matrix can be solved such that

$$AA^{-1} = AA^{-1} = I$$

... every non-singular matrix can also be transformed into a vector of latent roots λ (with elements λ_i) by solving the 'characteristic equation' using polynomial expansion.

$$\lambda V = AV$$

$$\lambda V - AV = (\lambda I - A)V = 0$$

Solution for V exists if $\det(\lambda I - A) = 0$

There are *n characteristic roots* to this equation

$$(\lambda_i I - A)v_i = 0$$

v_i is the *characteristic vector* for root i.

Eigenanalysis

Many covariance matrices can be mathematically decomposed into a product $C = VLV^{-1}$

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \begin{bmatrix} v_{1,1} & v_{1,2} & v_{1,3} \\ v_{2,1} & v_{2,2} & v_{2,3} \\ v_{3,1} & v_{3,2} & v_{3,3} \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix}^T \begin{bmatrix} v_{1,1} & v_{1,2} & v_{1,3} \\ v_{2,1} & v_{2,2} & v_{2,3} \\ v_{3,1} & v_{3,2} & v_{3,3} \end{bmatrix}$$

λ_i are *eigenvalues* v_i are *eigenvectors*

Each eigenvalue has its associated eigenvector.

Customary to arrange eigenvalues in descending order such that the largest one is first, etc.

Eigenvalues and Eigenvectors

Eigenvectors are scaled so that V is an orthogonal matrix.
Eigenvalues will all be > 0. $tr(A) = \sum_i \lambda$

Decomposition of a matrix into eigenvalues and eigenvectors rearranges dimensions in an n dimensional space such that the axes are all perpendicular.

Eigenvalues tell us the proportion of total variability in a matrix associated with its corresponding eigenvector.
i.e. eigenvector corresponding to largest eigenvalue tell us the dimension (axis) that generates the maximum amount of individual variability in the variables.

Next Time. . .