Mathematics Cheat Sheet for Population Biology

James Holland Jones,
Department of Anthropological Sciences
Stanford University

August 5, 2006

1 Introduction

If you fake it long enough, there comes a point where you aren't faking it any more. Here are some tips to help you along the way...

2 Calculus

Derivative The definition of a derivative is as follows. For some function $f(x)$,

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

2.1 Differentiation Rules

It is useful to remember the following rules for differentiation. Let $f(x)$ and $g(x)$ be two functions

2.1.1 Linearity

$$\frac{d}{dx}\left(af(x) + bg(x)\right) = af'(x) + bg'(x)$$

for constants a and b.

2.1.2 Product rule

$$\frac{d}{dx}\left(f(x)g(x)\right) = f'(x)g(x) + f(x)g'(x)$$

2.1.3 Chain rule

$$\frac{d}{dx}g(f(x)) = g'(f(x))f'(x)$$

2.1.4 Quotient Rule

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$

2.1.5 Some Basic Derivatives

$$\frac{d}{dx}x^a = ax^{a-1}$$

$$\frac{d}{dx}\frac{1}{x^a} = -\frac{a}{x^{a+1}}$$

$$\frac{d}{dx}e^x = e^x$$

$$\frac{d}{dx}a^x = a^x \log a$$

$$\frac{d}{dx}\log|x| = \frac{1}{x}$$

2.1.6 Convexity and Concavity

It is very easy to get confused about the convexity and concavity of a function. The technical mathematical definition is actually somewhat at odds with the colloquial usage. Let $f(x)$ be a twice differentiable function in an interval I. Then:

$$\begin{aligned} f''(x) \geq 0 \quad &\Rightarrow \quad f(x) \ \textit{convex} \\ f''(x) \leq 0 \quad &\Rightarrow \quad f(x) \ \textit{concave} \end{aligned} \tag{1}$$

If you think about a profit function as a function of time, a convex function would show increasing marginal returns, while a concave function would show decreasing marginal returns.

This leads into an important theorem (particularly for stochastic demography), known as Jensen's Inequality. For a convex function $f(x)$,

$$\mathbb{E}\left[f(X)\right] \geq f(\mathbb{E}\left[X\right]).$$

2.2 Taylor Series

$$T(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

where $f^{(k)}(a)$ denotes the kth derivative of f evaluated at a, and $k! = k(k-1)(k-2)\ldots(1)$.

For example, we can approximate e^r at $a = 0$:

$$e^r \approx 1 + r + \frac{r^2}{2} + \frac{r^3}{6} \ldots$$

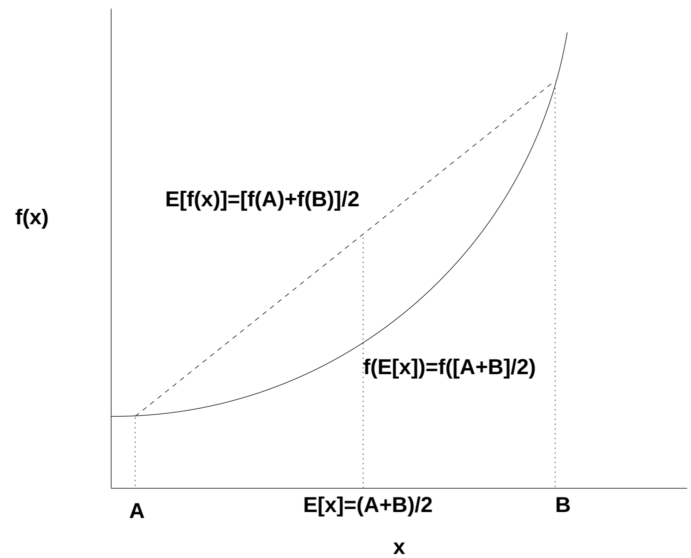

Figure 1: Illustration of Jensen's Inequality.

2.3 Jacobian

For a system of equations, $F(x)$ and $G(\lambda)$, the Jacobian matrix is

$$\mathbf{J} = \begin{pmatrix} \partial F/\partial x & \partial F/\partial \lambda \\ \partial G/\partial x & \partial G/\partial \lambda \end{pmatrix}.$$

This is very important for the analysis of stability of interacting models such as those for epidemics and predator-prey systems. The equilibrium of a system is stable if and only if the real parts of all the eigenvalues of $\mathbf{J}$ are negative.

2.4 Integration

Linearity

$$\int [af(x) + bg(x)]\, dx = a \int f(x)dx + b \int g(x)dx$$

Integration by Parts

$$\int u \cdot v'\, dx = u \cdot v - \int v \cdot u'\, dx$$

Some Useful Facts About Integrals

$$\int \frac{f'(x)}{f(x)} dx = \log |f(x)|$$

$$\int x^a dx = \frac{x^{a+1}}{a+1}, \quad a \neq -1$$

$$\int e^x dx = e^x$$

$$\int \frac{dx}{x} = \log|x|$$

2.5 Definite Integrals

$$\int_a^b f(x)dx = [F(x)]_a^b = F(b) - F(a)$$

2.5.1 Expectation

For a continuous random variable X with probability density function $f(x)$, the expected value is

$$\mathbb{E}(X) = \int_\Omega x f(x) dx$$

where the integral is taken over the set of all possible outcomes Ω.

For example, the average age of mothers of newborns in a stable population:

$$A_B = \int_\alpha^\beta a e^{-ra} l(a) m(a) da$$

Since (from the Euler-Lotka equation) the probability that a mother will be a years old in a stable population is $f(a) = e^{-ra} l(a) m(a)$.

Some Properties of Expectation

$$\mathbb{E}[aX] = a\mathbb{E}[X]$$

For two discrete random variables, X and Y,

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

2.5.2 Variance

For a continuous random variable X with probability density function $f(x)$ and expected value μ, the variance is

$$\mathbb{V}(X) = \int_\Omega (x - \mu)^2 f(x) dx$$

A useful formula for calculating variances:

$$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

2.6 Exponents and Logarithms

Properties of Exponentials

$$x^a x^b = x^{a+b}$$

$$\frac{x^a}{x^b} = x^{a-b}$$

$$x^a = e^{a \log x}$$

Complex Case

$$e^z = e^{a+bi} = e^a e^{bi} = e^a(\cos b + i \sin b)$$

$$(x^a)^b = x^{ab}$$

$$x^{-a} = \frac{1}{x^a}$$

The logarithm to the base e, where e is defined as

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

Assume that $\log \equiv \log_e$. Logarithms to other bases will be marked as such. For example: $\log_{10}$, $\log_2$, etc.

This is an important for demography:

$$\lim_{n \to \infty} \left(1 + \frac{r}{n}\right)^n = e^r$$

Properties of Logarithms

$$\log x^a = a \log x$$

$$\log ab = \log a + \log b$$

$$\log \frac{a}{b} = \log a - \log b$$

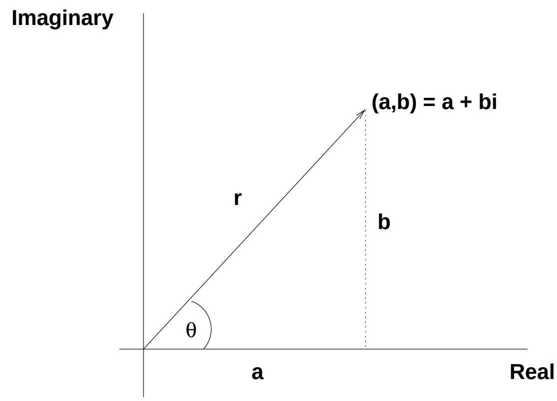

Figure 2: Argand diagram representing a complex number $z = a + bi$.

Complex Numbers We encounter complex numbers frequently when we calculate the eigenvalues of projection matrices, so it is useful to know something about them. Imaginary number: $i = \sqrt{-1}$. Complex number: $z = a + bi$, where a is the real part and b is a coefficient on the imaginary part.

It is useful to represent imaginary numbers in their polar form. Define axes where the abscissa represents the real part of a complex number and the ordinate represents the imaginary part (these axes are known as an *Argand diagram*). This vector, $a + bi$ can be represented by the angle θ and the radius of the vector rooted at the origin to point (a, b). Using trigonometric definitions, $a = r \sin\theta$ and $b = r\cos\theta$, we see that

$$z = a + ib = r(\cos\theta + i\sin\theta).$$

Believe it or not, this comes in handy when we interpret the transient dynamics of a population.

Let z be a complex number with real part a and imaginary part b,

$$z = a + bi$$

Then the complex conjugate of z is

$$\bar{z} = a - bi$$

Non-real eigenvalues of demographic projection matrices come in conjugate pairs.

3 Linear Algebra

A matrix is a rectangular array of numbers

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

A vector is simply a list of numbers

$$\mathbf{n}(t) = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$$

A scalar is a single number: $\lambda = 1.05$

We refer to individual matrix elements by indexing them by their row and column positions. A matrix is typically named by a capital (bold) letter (e.g., $\mathbf{A}$). An element of matrix $\mathbf{A}$ is given by a lowercase a subscripted with its indices. These indices are subscripted following the the lowercase letter, first by row, then by column. For example, a_{21} is the element of $\mathbf{A}$ which is in the second row and first column.

Matrix Multiplication

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \begin{bmatrix} a_{11}n_1 + a_{12}n_2 \\ a_{21}n_1 + a_{22}n_2 \end{bmatrix}$$

Multiply each row element-wise by the column

For Example,

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 6 \\ 7 \end{bmatrix} = \begin{bmatrix} (2 \cdot 6) + (3 \cdot 7) \\ (4 \cdot 6) + (5 \cdot 7) \end{bmatrix} = \begin{bmatrix} 33 \\ 59 \end{bmatrix}$$

Matrix Addition or Subtraction

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

Multiplying a Matrix by a Scalar

$$\lambda \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{bmatrix}$$

$$4 \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 8 & 12 \\ 16 & 20 \end{bmatrix}$$

Systems of Equations Matrix notation was invented to make solving simultaneous equations easier.

$$\begin{aligned} y_1 &= ax_1 + bx_2 \\ y_2 &= cx_1 + dx_2 \end{aligned}$$

In matrix notation:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

3.1 Eigenvalues and Eigenvectors

A scalar λ is an eigenvalue of a square matrix $\mathbf{A}$ and $\mathbf{w} \neq \mathbf{0}$ is its associated eigenvector if

$$\mathbf{A}\mathbf{w} = \lambda\mathbf{w}.$$

Eigenvalues of $\mathbf{A}$ are calculated as the roots of the characteristic equation,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0,$$

where $\mathbf{I}$ is the identity matrix, a square matrix with ones along the diagonal and zeros elsewhere.

For example, we can calculate the eigenvalues for the matrix,

$$\mathbf{A} = \begin{bmatrix} f_1 & f_2 \\ p_1 & 0 \end{bmatrix}.$$

Solve the characteristic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$:

$$(\mathbf{A} - \lambda\mathbf{I}) = \begin{bmatrix} f_1 & f_2 \\ p_1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} f_1 - \lambda & f_2 \\ p_1 & -\lambda \end{bmatrix}$$

$$\det(\mathbf{A} - \lambda\mathbf{I}) = -(f_1 - \lambda)\lambda - f_2 p_1$$

$$\lambda^2 - f_1\lambda - f_2 p_1 = 0$$

Use the quadratic equation to solve for λ:

$$\frac{-f_1 \pm \sqrt{f_1^2 - 4f_2 p_1}}{2f_1}$$

Numerical Example Define:

$$\mathbf{A} = \begin{bmatrix} 1.5 & 2 \\ 0.5 & 0 \end{bmatrix} \tag{2}$$

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{bmatrix} 1.5 - \lambda & 2 \\ 0.5 & -\lambda \end{bmatrix}$$

$$\lambda^2 - 1.5\lambda - 1 = 0$$

$$(\lambda - 2)(\lambda + 0.5) = 0$$

The roots of this are $\lambda = 2$ and $\lambda = -0.5$. A $k \times k$ matrix will have k eigenvalues. If a matrix is non-negative, irreducible, and primitive, one of these eigenvalues is guaranteed to be real, positive, and strictly greater than all the others.

Analytic Formula for Eigenvalues: The 2×2 Case

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The eigenvalues are:

$$\lambda_{\pm} = T/2 \pm \sqrt{(T/2)^2 - D}$$

where $T = a + d$ is the trace and $D = ad - bc$ is the determinant of matrix $\mathbf{A}$.