Appendix 1: Notes on the algebra of the generalized logarithmic function

Mark Carey, John Boland, Gunnar Keppel

November 27, 2022

1 The relation of the generalised logarithm used in this study to previous versions in the literature

The one-parameter Box-Cox transformations are defined as

$$y_i^{(\lambda)} = \begin{cases} \dfrac{y_i^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \ln y_i & \text{if } \lambda = 0, \end{cases}$$

Tsalis (2016) gives an integral definition of this:

$$\ln_w (x) = \int_1^x \frac{dt}{t^w} \tag{1}$$

If we let the bottom integrand in the integral definition of Equation 1 slide we can instead construct the function as

$$\ln_q (x) = \int_q^x \frac{dt}{t^q} \tag{2}$$

which evaluates, for $q \neq 1$ to

$$\ln_q x = \frac{x^p - q^p}{p} \tag{3}$$

where $p = 1 - q$ and $\ln_q x = \ln x$ for $q = 1$. This function smoothly transforms between an identity ($x = x$) transform and a logarithmic transform in a way in which the Box Cox transform and other generalised logarithmic transformations (Box & Cox, 1964 ; Martinez, 2008) don't (they go from $y = x + 1$ to $y = \ln(x)$).

2 The algebra of the generalised logarithmic function

It is important too keep in mind that the algebra of partial logarithms (i.e., $\ln_q x$ for $q \in (0, 1)$) are not the same as for the regular logarithm ($q = 1$). Importantly the identities

$$\ln a^b = b \ln a$$

and

$$\ln ab = \ln a + \ln b$$

only hold for $\ln_q$ when $q = 1$ i.e., where you are dealing with the regular logarithm.

2.1 The inverse function of $\ln_q(x)$: the generalised exponential

If we have $q \in [0, 1)$ and $p = 1 - q$ and $\ln_q(x)$ then defined as

$$\ln_q x = \frac{x^p - q^p}{p} \tag{4}$$

Now, in general we have

$$y = e^x \iff \ln y = x$$

so, for the case we are interested in we have

$$y = \sqrt[p]{px + q^p} = (px + q^p)^{1/p}. \tag{5}$$

Thus we have as our expression for $\exp_q$

$$\exp_q x = (px + q^p)^{1/p}. \tag{6}$$

This dovetails with the asymptotic definition of the exponential function, as we can see if we further define $n = 1/p$ so $n \to \infty$ as $p \to 0$ i.e.,

$$\exp_q x = \left(\frac{x}{n} + q^p\right) n \tag{7}$$

and as $q \to 1$, $p \to 0$ and we recover:

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

2.2 Partitioning the logarithmic function into partial logarithms

If we define a function like that given in Equation 4 (for our case we have $\phi = cA^z$), which can be conceptually thought of as a "partial logarithm", then a natural question is how the logarithm can be partitioned i.e., is there a $\bar{q}$ such that

$$\ln_{\bar{q}}(\ln_q(\phi)) = \ln \phi? \tag{8}$$

The answer is no.

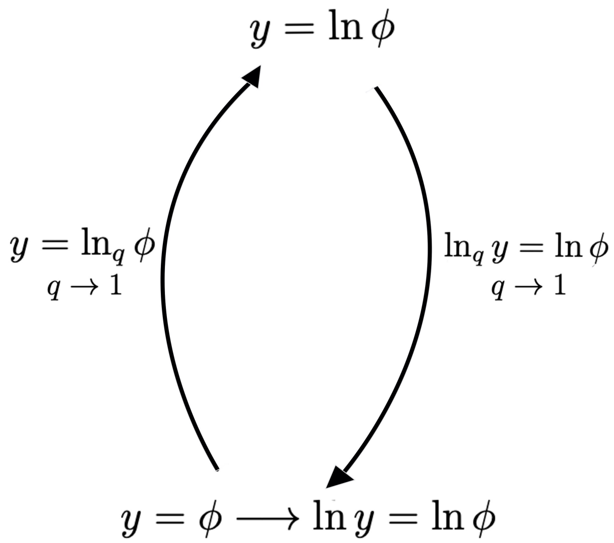
If the answer was yes, this would imply that one could "log both sides" of an equation in the same way that one could do with the regular $\ln \phi$ function, which would be nice, but which does not work.

If we have complementary values of q ie: $\exists q$ and $\bar{q}$ that for all $x > 1$ then, if the construction proposed in Equation 8 were possible, then that would imply that q and $\bar{q}$ were related via a solution of the equation

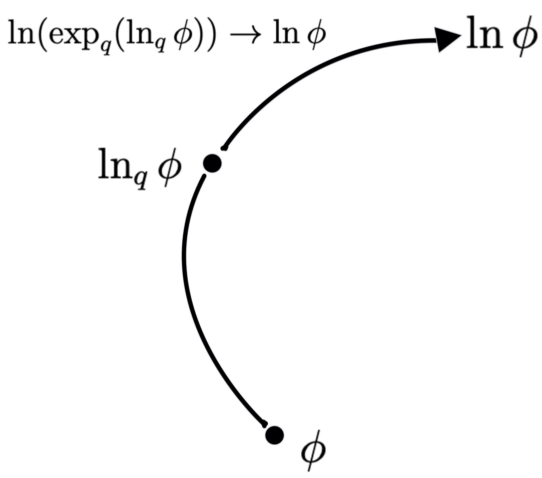
$$\frac{\phi^{\bar{p}} - \bar{q}^{\bar{p}}}{\bar{p}} = (p \ln \phi + q^p)^{1/p}. \tag{9}$$

If one performs a a Taylor series expansion around $x = 1$ on both sides of Equation 9 however, one can convince oneself rather rapidly that the RHS and LHS of Equation 9 are different functions and hence that this is not an equation at all[1].

Conceptually this is a consequence of the fact that tracing a path though function space with intermediate models is not the same if one partially logs the response variable versus the function equated to the response variable. This is illustrated in Figure 1).



(a)



(b)

Figure 1: (a) distinct paths in function space traced by intermediate models by two different formulations of the generalized power law using the $\ln_q$ function. (b) the order of operations required to trace a single path in function space when partitioning the logarithm. For our application we have y being species richness ($y = S$) and $\phi = cA^z$, where A is either island area or sample plot area, so that $\ln \phi = \ln c + z \ln A$

[1]At this point I should credit Timothy Nguyen, who during a series of conversations managed to help clarify some points about these functions not originally appreciated by the authors.

Instead there is a complementary function which takes the function in question from $\ln \phi$ to $\ln_q \phi$ along the same trajectory in function space. This complementary function will be:

$$\exp_q(\ln \phi) = (p \ln \phi + q^p)^{\frac{1}{p}} \tag{10}$$

i.e., the function which traces the other way from $\ln x$ in the function space. Instead of going from $y = x$ to $\ln_q x$ you go from $\ln x$ to $\ln_q x$.

So if we have

$$S = \ln_q(cA^z) \tag{11}$$

the equivalent model to solve is actually

$$\ln(\exp_q(S)) = \ln c + z \ln A \tag{12}$$

and the equivalent model for

$$\ln_q(S) = \ln c + z \ln A \tag{13}$$

So the model, termed the SqA model in the main text is equivalent to

$$S = \exp_q(\ln(cA^z)) \tag{14}$$

Each set of models traces a different path in the function space and forms a disjoint set, although this should not substantially affect the statistics of my results. The correct formulation for $\ln_{\bar{q}}$ is thus

$$\ln_{\bar{q}} x = \ln_{\bar{q}}(\exp_q(\ln_q(x))), \tag{15}$$

where the order of operations is necessary to keep in mind.

Basically, what the above suggests is that the most efficient and possibly the *only* way which preserves the path in function space is to retrace your steps to linear space and then shoot for your new destination (i.e., the new partial logarithmic function) along that path.

2.3 Relation to hybrid algebraic operators

Historically, the reason to introduce logarithms, first by Napier, was to transform large multiplications into simpler additions to facilitate the simplification of difficult problems. It is not surprising then that the introduction of a generalised logarithmic function also has implications for the inter-transformation of basic algebraic operators. Most immediately it introduces an intermediate operator which is part way between addition and multiplication which I here designate as $\oplus$.

$$a \oplus_q b = a \times b \text{ for } q = 1 \tag{16}$$
$$= a + b \text{ for } q = 0 \tag{17}$$

where $q \in (0, 1)$

So if we have $4 \oplus 5$ this will be somewhere between 9 and 20 depending on q. More rigorously, we can define the $\oplus$ operator as

$$a \oplus_q b = \exp_q (\ln_q a + \ln_q b). \tag{18}$$

This has the qualities we want, in that is equal to $a + b$ when $q = 0$ (both ln and the exp transforms turned off) and $a \times b$ when $q = 1$ ie:

$$\exp(\ln a + \ln b) = \exp(\ln(ab)) = ab$$

$$a \oplus_q b = \left(p \left(\frac{a^p - q^p}{p} + \frac{b^p - q^p}{p} \right) + q^p \right)^{1/p} \tag{19}$$

which simplifies to

$$a \oplus_q b = (a^p + b^p - q^p)^{1/p} \tag{20}$$

for $q = 0$ $(p = 1)$ it is easy to see this is $a + b$. It is also clear

$$\lim_{q \to 1} (a^p + b^p - q^p)^n = ab \tag{21}$$

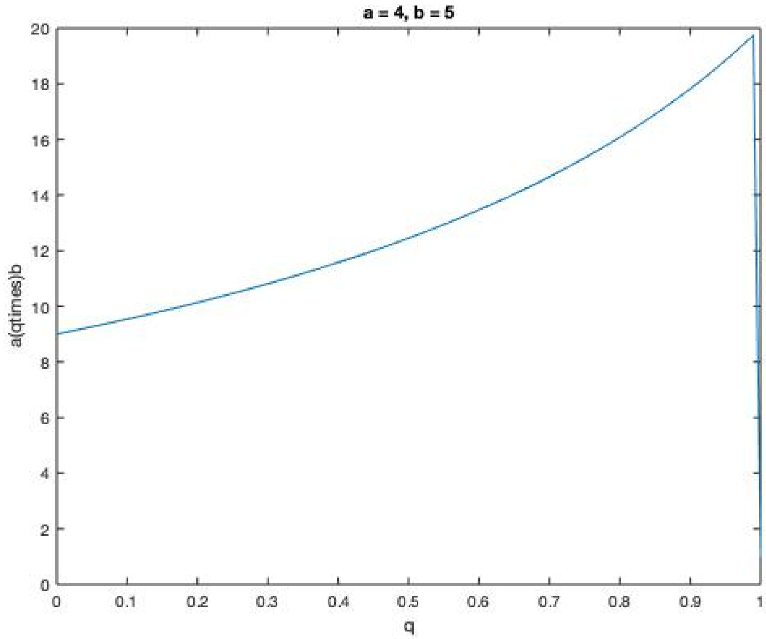


Figure 2: Here we have, given the definition of 20, a graph of $a \oplus_q$ for $a = 4$, $b = 5$. Ignore the very far right section of the graph as this is the region, given the resolution of the sampling, where the formula used to generate the graphic breaks down.

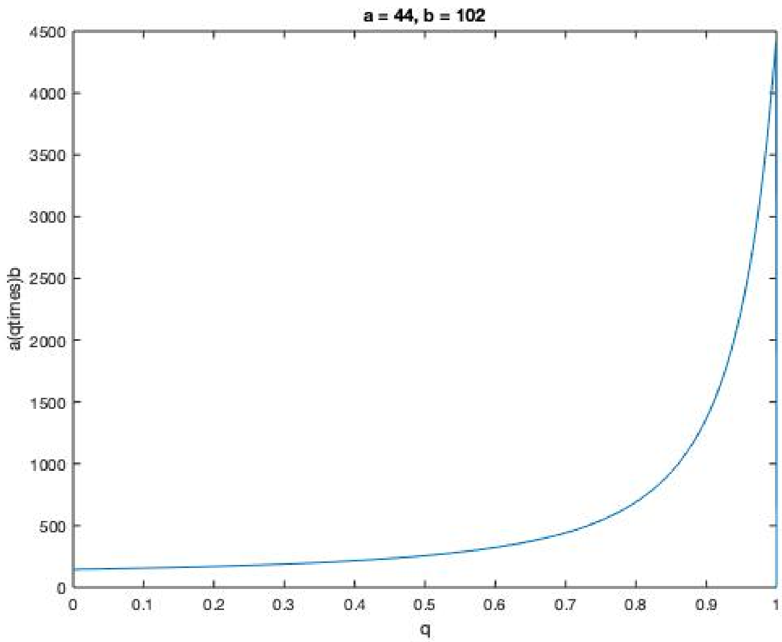


Figure 3: $a \oplus_q$ b for $a = 44$, $b = 102$

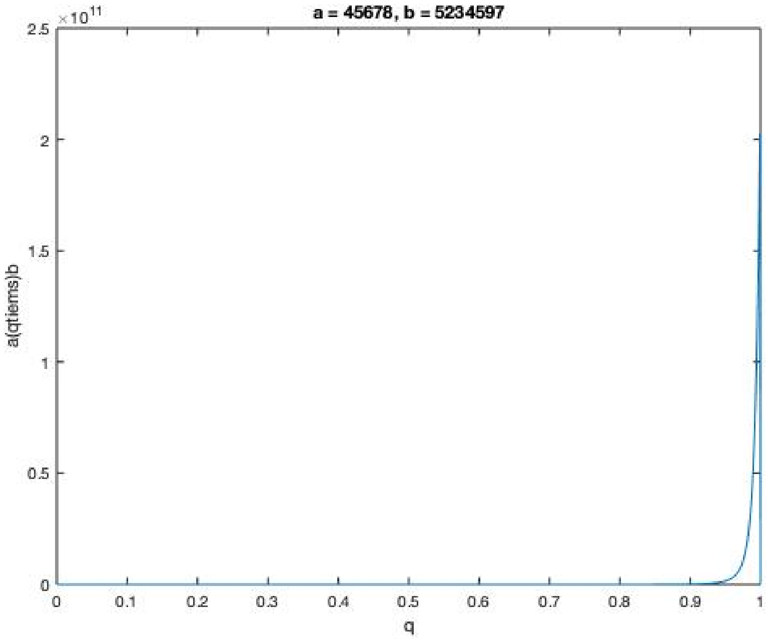


Figure 4: $a \oplus_q$ for $a = 45678$, $b = 5234597$

We can notice, comparing Figures 2, 3 and 4, that the operator is rather "lop-sided" as far as where the multiplication dominates over the addition i.e., the transition from additive to multiplicative behaviour of $\oplus$ is pushed closer to $q \approx 1$ as we increase the magnitude of the integers considered.

References

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243.

Martinez, A. S., González, R. S., & Terçariol, C. A. S. (2008). Generalized exponential function and some of its applications to complex systems. *arXiv preprint arXiv:0812.3071*.

Roegel, D. (2010). Napier's ideal construction of the logarithms.

Tsallis, C. (2016). Inter-occurrence times and universal laws in finance, earthquakes and genomes. *Chaos, Solitons & Fractals*, *88*, 254–266.