

SSlogis {stats}

R Documentation

Self-Starting Nls Logistic Model

Description

This `selfstart` model evaluates the logistic function and its gradient. It has an `initial` attribute that creates initial estimates of the parameters `Asym`, `xmid`, and `scal`.

Usage

```
SSlogis(input, Asym, xmid, scal)
```

Arguments

`input` a numeric vector of values at which to evaluate the model.

`Asym` a numeric parameter representing the asymptote.

`xmid` a numeric parameter representing the `x` value at the inflection point of the curve. The value of `SSlogis` will be `Asym/2` at `xmid`.

`scal` a numeric scale parameter on the `input` axis.

Value

a numeric vector of the same length as `input`. It is the value of the expression `Asym/(1+exp((xmid-input)/scal))`. If all of the arguments `Asym`, `xmid`, and `scal` are names of objects the gradient matrix with respect to these names is attached as an attribute named `gradient`.

Author(s)

José Pinheiro and Douglas Bates

See Also

[nls](#), [selfStart](#)

Examples

```
Chick.1 <- ChickWeight[ChickWeight$Chick == 1, ]
SSlogis(Chick.1$Time, 368, 14, 6) # response only
Asym <- 368; xmid <- 14; scal <- 6
SSlogis(Chick.1$Time, Asym, xmid, scal) # response and gradient
getInitial(weight ~ SSlogis(Time, Asym, xmid, scal), data = Chick.1)
## Initial values are in fact the converged values
fm1 <- nls(weight ~ SSlogis(Time, Asym, xmid, scal), data = Chick.1)
summary(fm1)
```

8.9 Balanced longitudinal data - Random coefficients and cubic smoothing splines

This section illustrates the use of random coefficients and cubic smoothing splines for the analysis of balanced longitudinal data.

The implementation of cubic smoothing splines in `asreml()` is based on the mixed model formulation of Verbyla et al. [1999]. More recently the methodology has been extended so that the user can specify knot points; in the original approach the knot points were taken to be the ordered set of unique values of the explanatory variable. The specification of knot points is particularly useful if the number of unique values in the explanatory variable is large, or if units are measured at different times.

These data were originally reported by Draper and Smith [1998, ex24N, p559] and have recently been reanalysed by Pinheiro and Bates [2000, p338]. The data are trunk circumferences (in millimetres) of each of 5 trees taken at 7 times (Figure 8.12). All trees were measured at the same time so that the data are balanced. The aim of the study is unclear, though both previous analyses involved modelling the overall *growth* curve, accounting for the obvious variation in both level and shape between trees.

Pinheiro and Bates [2000] used a nonlinear mixed effects modelling approach, in which they modelled the growth curves by a three parameter logistic function of age:

$$y = \frac{\phi_1}{1 + \exp[-(x - \phi_2)/\phi_3]} \quad (8.13)$$

where y is the trunk circumference, x is the tree age in days since December 31 1968, ϕ_1 is the asymptotic height, ϕ_2 is the inflection point or the time at which the tree reaches $0.5\phi_1$, ϕ_3 is the time elapsed between trees reaching half and about $3/4$ of ϕ_1 .

The data frame `orange` contains:

```
> orange <- asreml.read.table("orange.csv", header=T, sep=",")  
> names(orange)  
[1] "Tree"  "x"    "circ"  "Season"
```

where `Tree` is a factor with 5 levels, `x` is tree age in days since 31 December 1968, `circ` is the trunk circumference and `Season` is a factor with two levels, `Spring` and `Autumn`. The factor `Season` was included after noting that tree age spans several years and if converted to day of year, measurements were taken in either April/May (`Spring`) or September/October (`Autumn`).

Initially we restrict the dataset to tree 1 to demonstrate fitting cubic splines in `asreml()`. The model includes the intercept and linear regression of trunk circumference on x and an additional random term `spl(x)` which includes a random term with a special design matrix with $7 - 2 = 5$ columns which relate to the vector, δ whose elements $\delta_i, i = 2, \dots, 6$ are the second differentials of the cubic spline at the knot points. The second differentials of a natural cubic spline are zero at the first and last knot points [Green and Silverman, 1994].

```
> orange.asr <- asreml(circ ~ x, random = ~ spl(x),  
+ splinepoints = list(x = c(118,484,664,1004,1231,1372,1582)),  
+ data = orange, subset = Tree==1)
```

In this example the spline knot points are specifically given in the `splinepoints` argument. These extra points have no effect in this case as they are the seven ages existing in the data file. In this instance the analysis would be the same if the `splinepoints` argument was omitted.

```
> summary(orange.asr)$varcomp
```

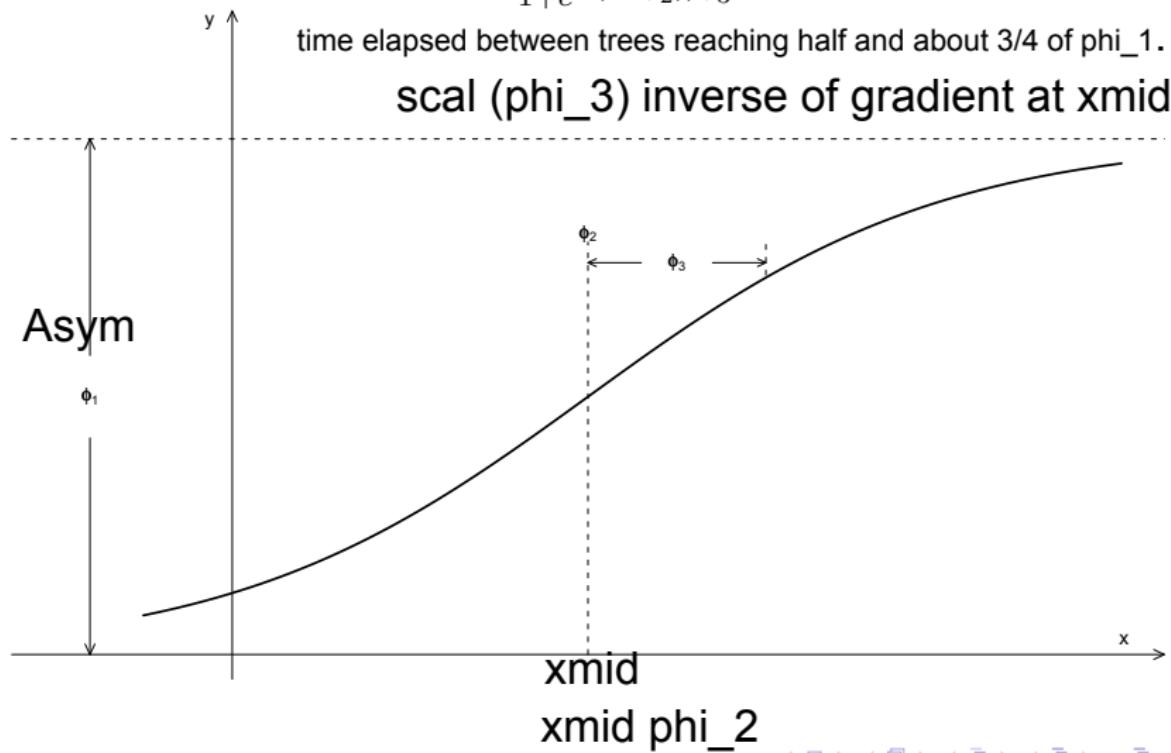
	gamma	component	std.error	z.ratio	constraint
<code>spl(x)</code>	0.07876884	3.954159	9.950608	0.3973786	Positive
<code>R!variance</code>	1.00000000	50.199529	37.886791	1.3249876	Positive

SSlogis(Time, Asym, xmid, scal)

The logistic growth model, SSlogis

$$y = \frac{\phi_1}{1+e^{-(x-\phi_2)/\phi_3}}$$

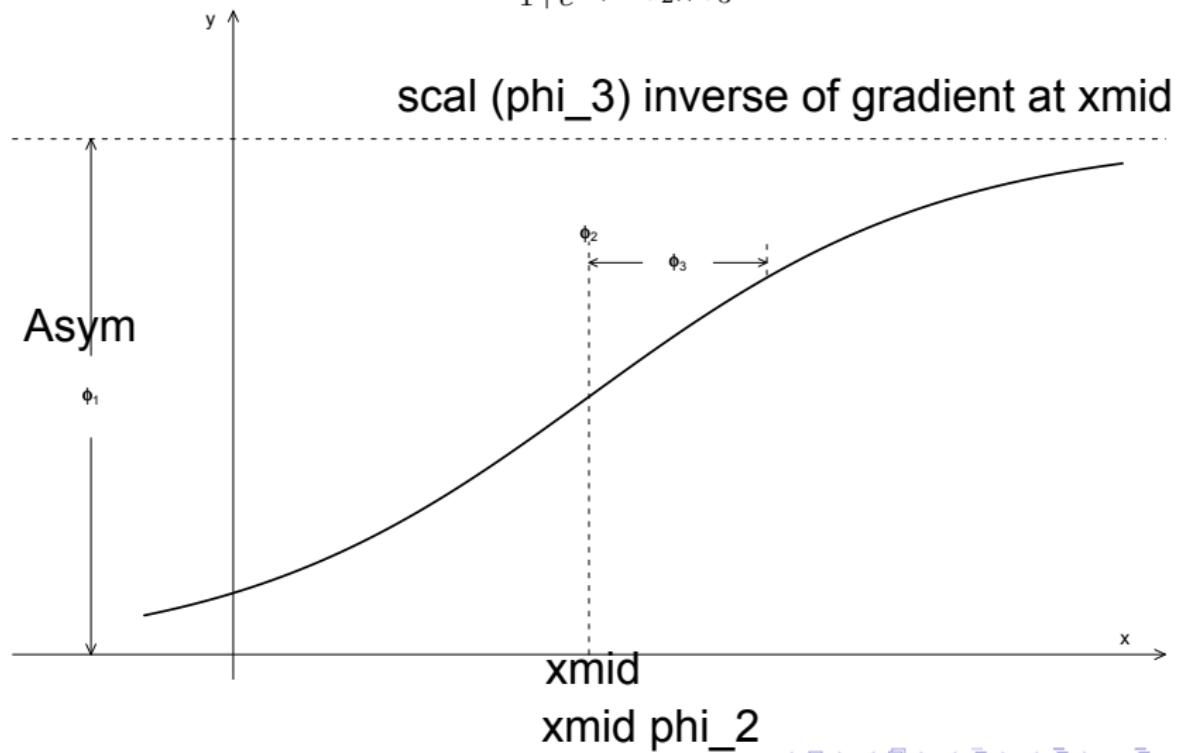
time elapsed between trees reaching half and about 3/4 of phi_1.
scal (phi_3) inverse of gradient at xmid



SSlogis(Time, Asym, xmid, scal)

The logistic growth model, SSlogis

$$y = \frac{\phi_1}{1+e^{-(x-\phi_2)/\phi_3}}$$



any theoretical considerations about the underlying mechanism producing the data.

Nonlinear models, on the other hand, are often *mechanistic*, i.e., based on a model for the mechanism producing the response. As a consequence, the model parameters in a nonlinear model generally have a natural physical interpretation. Even when derived empirically, nonlinear models usually incorporate known, theoretical characteristics of the data, such as asymptotes and monotonicity, and in these cases, can be considered as *semi-mechanistic* models. A nonlinear model generally uses fewer parameters than a competitor linear model, such as a polynomial, giving a more *parsimonious* description of the data. Nonlinear models also provide more reliable predictions for the response variable outside the observed range of the data than, say, polynomial models would.

To illustrate these differences between linear and nonlinear models, let us consider a simple example in which the expected height h_t of a tree at time t follows a three-parameter *logistic* growth model.

$$h_t = \phi_1 / \{1 + \exp [-(t - \phi_2) / \phi_3]\}. \quad (6.1)$$

As described in Appendix C.7, the parameters in (6.1) have a physical interpretation: ϕ_1 is the asymptotic height; ϕ_2 is the time at which the tree reaches half of its asymptotic height; and ϕ_3 is the time elapsed between the tree reaching half and $1/(1 + e^{-1}) \approx 3/4$ of its asymptotic height. The logistic model (6.1) is linear in one parameter, ϕ_1 , but nonlinear in ϕ_2 and ϕ_3 .

To make the example more concrete, suppose that $\phi_1 = 3$, $\phi_2 = 1$, and $\phi_3 = 1.2$ and that we initially want to model the tree growth for $0.4 \leq t \leq 1.6$. The logistic curve, shown as a solid line in Figure 6.1, is approximated very well in the interval $[0.4, 1.6]$ by the fifth-degree polynomial

$$h_t \approx -2.2911 + 16.591t - 44.411t^2 + 56.822t^3 - 31.514t^4 + 6.3028t^5$$

obtained as a least-squares fit to equally spaced t values in the interval $[0.4, 1.6]$. The polynomial fit, shown as a dashed line in Figure 6.1, is virtually indistinguishable from the logistic curve within this interval.

Unlike the coefficients in the logistic model, the coefficients in the polynomial approximation do not have any physical interpretation. Also, the linear polynomial model uses twice as many parameters as the logistic model to give comparable fitted values. Finally, the polynomial approximation is unreliable outside the interval $[0.4, 1.6]$. Figure 6.2, displaying the two curves over the extended interval $[0, 2]$, shows the dramatic differences between the curves outside the original range. We would expect growth curves to follow a pattern more like the logistic model than like the polynomial model.

Nonlinear mixed-effects models extend linear mixed-effects models by allowing the regression function to depend nonlinearly on fixed and random effects. Because of its greater flexibility, an NLME model is generally more

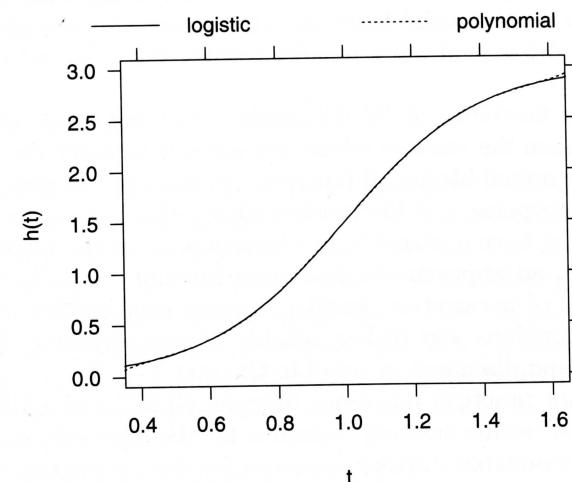


FIGURE 6.1. Logistic curve with parameters $\phi_1 = 3$, $\phi_2 = 1$, and $\phi_3 = 1.2$ and its fifth-order polynomial approximation over the interval $[0.4, 1.6]$.

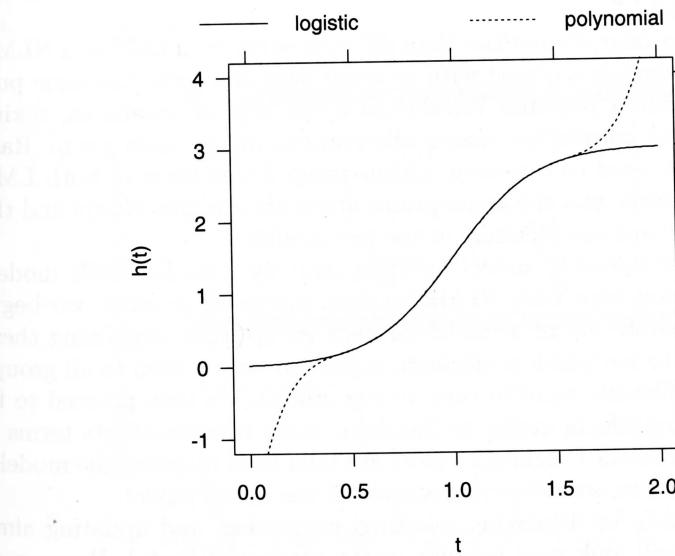


FIGURE 6.2. Logistic curve with parameters $\phi_1 = 3$, $\phi_2 = 1$, and $\phi_3 = 1.2$ and its fifth-order polynomial approximation over the interval $[0.4, 1.6]$, plotted over the interval $[0, 2]$.

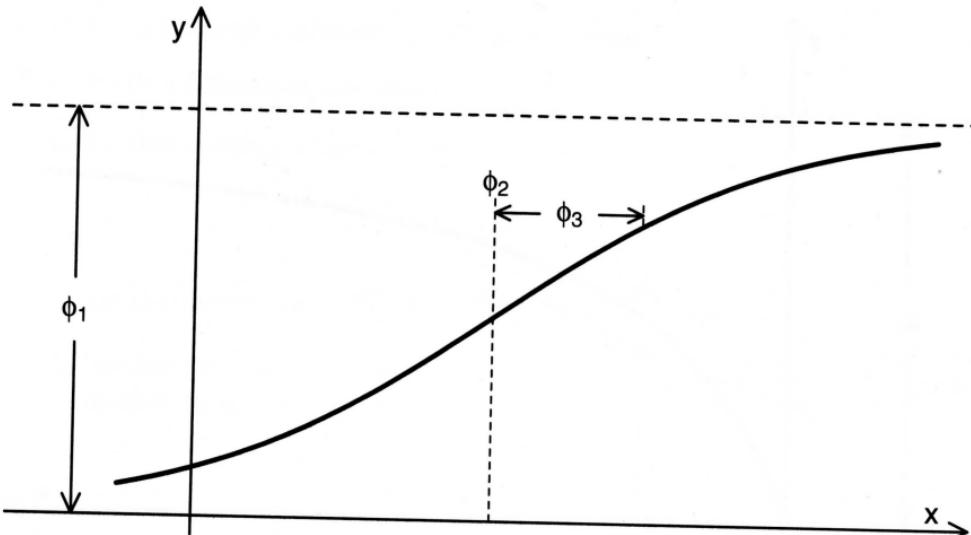


FIGURE C.7. The simple logistic model showing the parameters ϕ_1 , the horizontal asymptote as $x \rightarrow \infty$, ϕ_2 , the value of x for which $y = \phi_1/2$, and ϕ_3 , a scale parameter on the x -axis. If $\phi_3 < 0$ the curve will be monotone decreasing instead of monotone increasing and ϕ_1 will be the horizontal asymptote as $x \rightarrow -\infty$.

C.7 SSlogis—Simple Logistic Model

The simple logistic model is a special case of the four-parameter logistic model in which one of the horizontal asymptotes is zero. We write it as

$$y(x) = \frac{\phi_1}{1 + \exp [(\phi_2 - x)/\phi_3]}. \quad (\text{C.7})$$

For this model we do not require that the scale parameter ϕ_3 be positive. If $\phi_3 > 0$ then ϕ_1 is the horizontal asymptote as $x \rightarrow \infty$ and 0 is the horizontal asymptote as $x \rightarrow -\infty$. If $\phi_3 < 0$, these roles are reversed. The parameter ϕ_2 is the x value at which the response is $\phi_1/2$. It is the inflection point of the curve. The scale parameter ϕ_3 represents the distance on the x -axis between this inflection point and the point where the response is $\phi_1 / (1 + e^{-1}) \approx 0.73\phi_1$. These parameters are shown in Figure C.7.

C.7.1 Starting Estimates for SSlogis

The starting estimates are determined by:

1. Scaling and, if necessary, shifting the responses y so the transformed responses y' are strictly within the interval $(0, 1)$.
2. Taking the logistic transformation

$$z = \log[y'/(1 - y')]$$

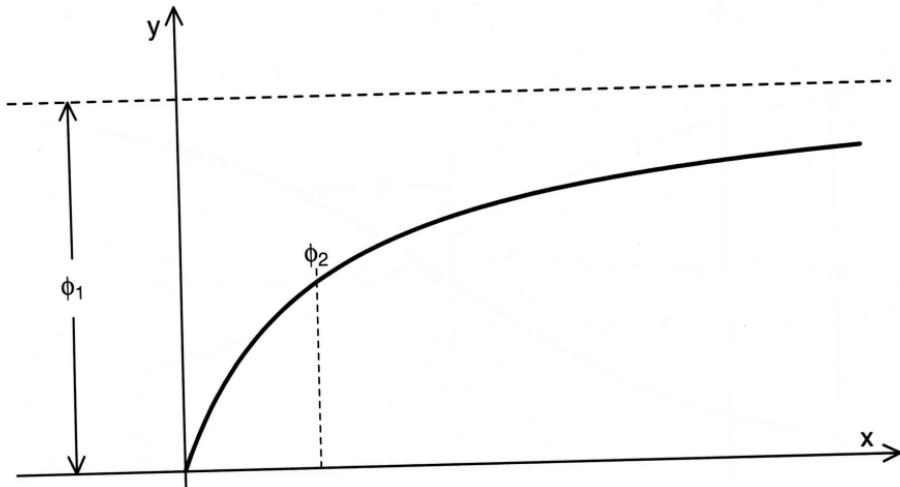


FIGURE C.8. The Michaelis–Menten model used in enzyme kinetics. The parameters are ϕ_1 , the horizontal asymptote as $x \rightarrow \infty$ and ϕ_2 , the value of x at which the response is $\phi_1/2$.

and fitting the simple linear regression model

$$x = a + bz.$$

3. Use $\phi_2^{(0)} = a$ and $\phi_3^{(0)} = b$ and an algorithm for partially linear models to fit

$$y = \frac{\phi_1}{1 + \exp[(\phi_2 - x)/\phi_3]}.$$

The resulting estimates are the final nonlinear regression estimates.

C.8 SS_{micmen}—Michaelis–Menten Model

The Michaelis–Menten model is used in enzyme kinetics to relate the initial rate of an enzymatic reaction to the concentration of the substrate. It is written

$$y(x) = \frac{\phi_1 x}{\phi_2 + x}, \quad (\text{C.8})$$

where ϕ_1 is the horizontal asymptote as $x \rightarrow \infty$ and ϕ_2 , the Michaelis parameter, is the value of x at which the response is $\phi_1/2$.

These parameters are shown in Figure C.8