

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
spl(x)	0.07876884	3.954159	9.950608	0.3973786	Positive
R!variance	1.00000000	50.199529	37.886791	1.3249876	Positive

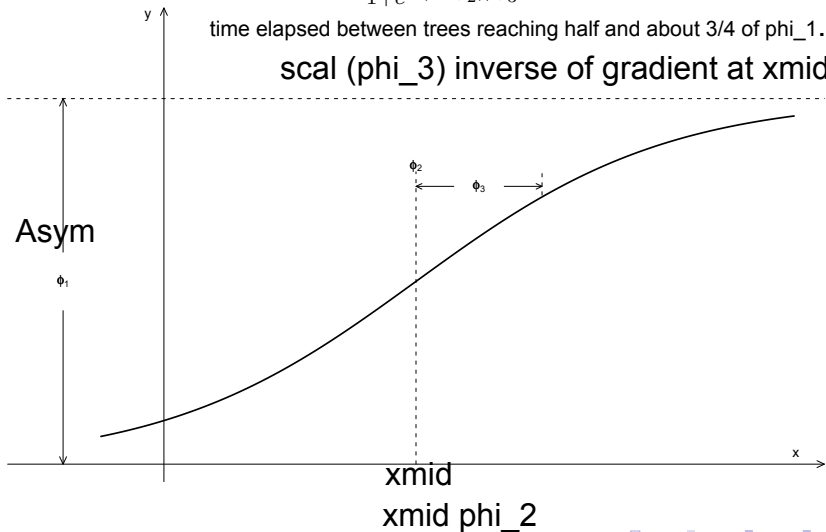
SSlogis(Time, Asym, xmid, scal)

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

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

