READING: Modern Applied Statistics with S, Venables & Ripley, Sections 8.1 & 8.2

Non-linear regression

ST552 Lecture 26

Charlotte Wickham

2019-03-11

Announcements

- **Extra** office hours:
  Fri (3/15) 1-2pm,
  Mon (3/18) 11am-12:30pm

In my office

Non-linear regression

$$y_i = \boxed{\overset{x_i^T\beta}{\eta(\mathbf{x_i}, \beta)}} + \epsilon_i \quad \epsilon_i \underset{i.i.d}{\sim} N(0, \sigma^2)$$

where η is a known function $\mathbf{x}_i$ is a vector of covariates and β a vector of p unknown parameters.

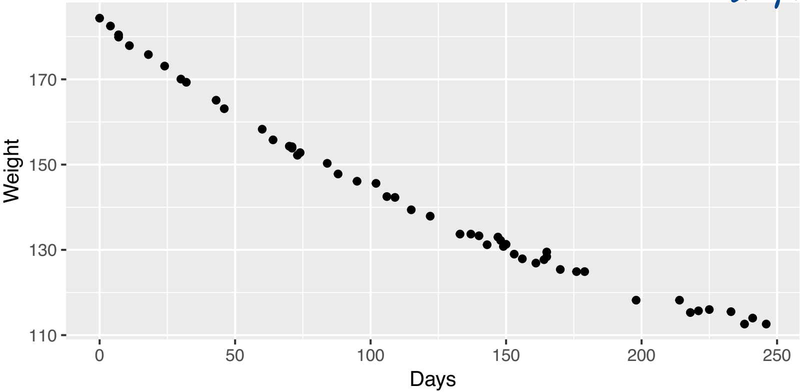
Distinctions:

- If $\eta(\mathbf{x}, \beta) = \mathbf{x}^T\beta$ then we are in the usual linear regression setting.
- If $\eta(\mathbf{x}, \beta)$ is unknown but we are willing to represent it using basis functions, we are in the smooth regression setting.

Example: `MASS::wtloss`

The data frame gives the weight, in kilograms, of an obese patient at 52 time points over an 8 month period of a weight rehabilitation programme.

$i = 1, \ldots, 52$ timepoints



Example: A potential model

Where did this come from?

$= \eta(t, (\beta_0, \beta_1, \theta)) + \varepsilon_t$

$$y_t = \beta_0 + \beta_1 2^{-t/\theta} + \epsilon_t$$

$\boxed{\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)}$

is this reasonable here?

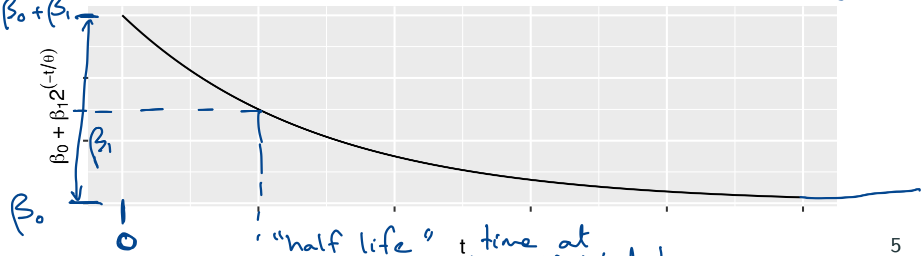
β_0, β_1 are linear parameters, θ is a non-linear parameter.

Your turn

- When $t = 0$, what is $\mathrm{E}(y_i)$? $\beta_0 + \beta_1$
- When $t = \infty$, what is $\mathrm{E}(y_i)$? β_0
- If $\mathrm{E}(y_i) = \beta_0 + \frac{1}{2}\beta_1$, what is t? θ

$\theta > 0$

$\frac{1}{2} = 2^{-t/\theta} \iff -\frac{t}{\theta} = -1$



"half life" time at which half total loss occurs

Fitting non-linear models

Under the Normal error assumption, the MLE of β, minimizes

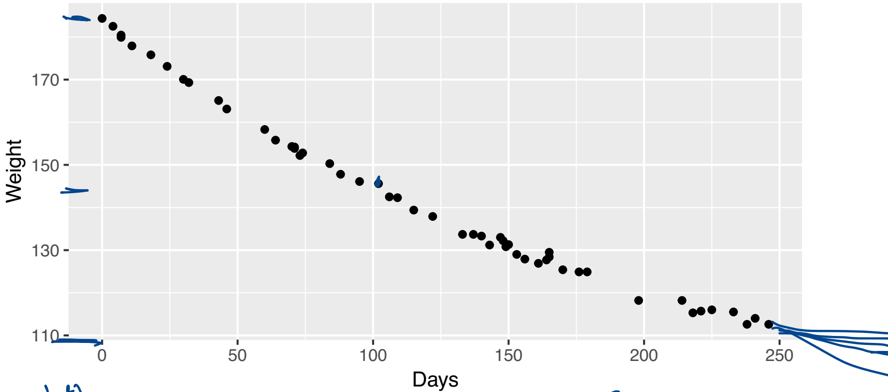
$$\sum_{i=1}^{n}(y_i - \underline{\eta(\mathbf{x_i}, \beta))^2}$$

(the sum of squared residuals) a.k.a non-linear least squares.

- There isn't in general a closed form solution, so iterative procedures are used.
  numerical
- This means you need to provide starting values from the parameters and check that the procedure converged.

Your turn

What might be reasonable values for β_0, β_1 and θ?



E(weight) @ time ∞

$\hat{\beta}_0 \approx 0$
≈ 110
≈ 100

$\hat{\beta}_1 \approx 75$
≈ 70
≈ 185

$\tilde{\theta} \approx 100$
≈ 90

In R: nls()

−8000 0 1

```
wtloss.st <- c(b0 = 100, b1 = 80, th = 100)
fit_nls <- nls(Weight ~ b0 + b1*2^(-Days/th),
  data = wtloss, start = wtloss.st)
fit_nls
```

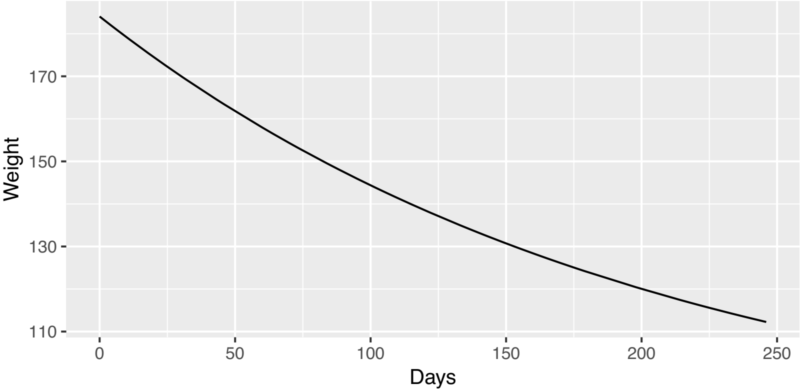
```
## Nonlinear regression model
##   model: Weight ~ b0 + b1 * 2^(-Days/th)
##    data: wtloss
##     b0     b1     th
##  81.37 102.68 141.91
##  residual sum-of-squares: 39.24
##
## Number of iterations to convergence: 4
## Achieved convergence tolerance: 5.259e-08
```

Check for convergence!

```r
summary(fit_nls)
```

```
##
## Formula: Weight ~ b0 + b1 * 2^(-Days/th)
##
## Parameters:
##    Estimate Std. Error t value Pr(>|t|)
## b0   81.374      2.269   35.86   <2e-16 ***
## b1  102.684      2.083   49.30   <2e-16 ***
## th  141.910      5.295   26.80   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8949 on 49 degrees of freedom
##
## Number of iterations to convergence: 4
## Achieved convergence tolerance: 5.259e-08
```

```
ggplot(wtloss, aes(Days, Weight)) +
  geom_line(aes(y = fitted(fit_nls))) +
geom_point( )
```



Example: `MASS::muscle`

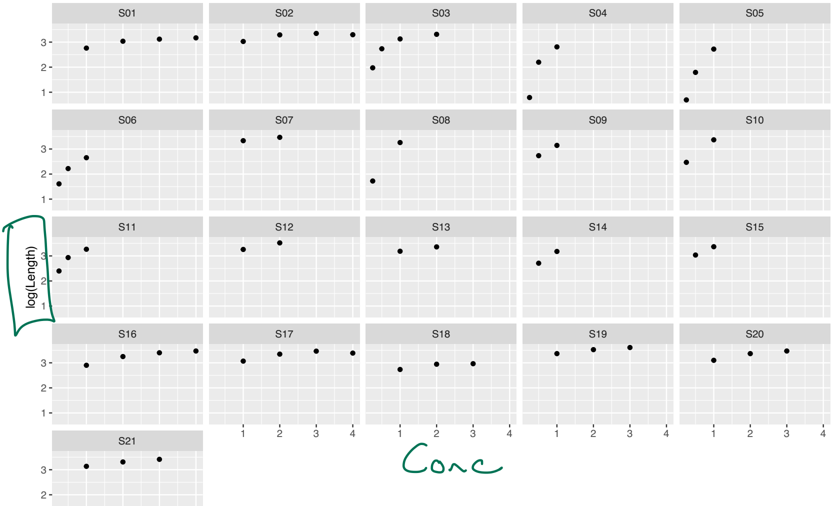
The purpose of this experiment was to assess the influence of calcium in solution on the contraction of heart muscle in rats. The left auricle of 21 rat hearts was isolated and on several occasions a constant-length strip of tissue was electrically stimulated and dipped into various concentrations of calcium chloride solution, after which the shortening of the strip was accurately measured as the response.

	Strip	Conc	Length
3	S01	1	15.8
4	S01	2	20.8
5	S01	3	22.6
6	S01	4	23.8
9	S02	1	20.6
10	S02	2	26.8

first six rows

Example: `MASS::muscle`

```
ggplot(muscle, aes(Conc, log(Length))) +
  geom_point() +
  facet_wrap(~ Strip)
```



Example: The same parameters for every strip

length Conc

$$\log \overset{\text{length}}{y_{ij}} = \underbrace{\alpha}_{\text{linear}} + \underbrace{\beta}_{\text{linear}}\underbrace{\rho}_{\text{non-linear}}^{x_{ij}} + \epsilon_{ij}$$

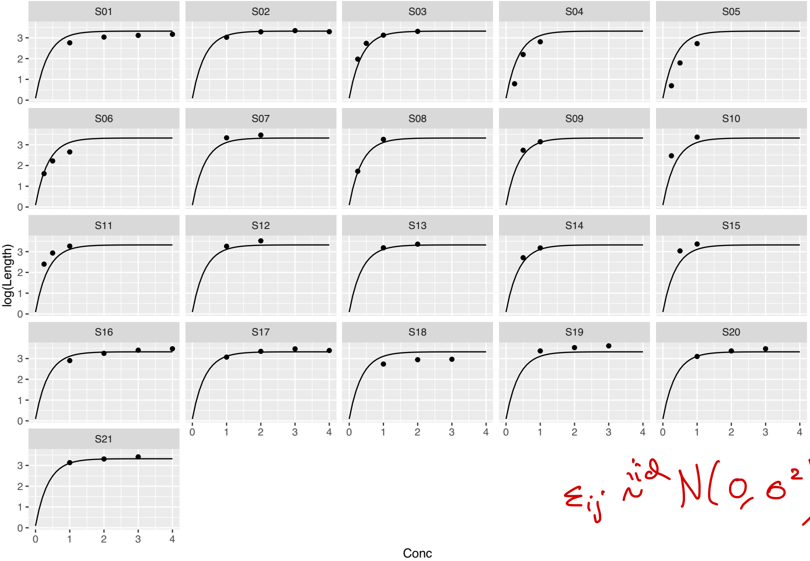
j indexes Strip, i the ith measurement on a strip.

```
fit_all <- nls(log(Length) ~ cbind(1, rho^Conc), muscle,
  start = list(rho = 0.05), algorithm = "plinear")

pred_grid <- with(muscle, expand.grid(Conc = seq(0, 4, 0.1),
  Strip = unique(Strip)))
pred_grid$fit_all <- predict(fit_all, pred_grid)
```

The `plinear` algorithm takes advantage of the linear parameters.

Example: The same parameters for every strip



$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$

Example: Different parameters for every strip

same for all rats

$$\log y_{ij} = \boxed{\alpha_j} + \boxed{\beta_j} \rho^{x_{ij}} + \epsilon_{ij}$$

```
fit_each <- nls(log(Length) ~ alpha[Strip] + beta[Strip] * rho^Conc,
  muscle,
  start = list(rho = coef(fit_all)[1], alpha = rep(coef(fit_all)[2], 21),
    beta = rep(coef(fit_all)[3], 21)))

pred_grid$fit_each <- predict(fit_each, pred_grid)
```

Example: Different parameters for every strip