Introduction to **Nonlinear Regression**

Evgenia Ageeva

26.04.2010

Motivation:

So far: linear mixed-effects models

$\longrightarrow$

Before mixed models,

lets look at nonlinear regression

$\longrightarrow$

Future: nonlinear mixed-effects models

Motivation:

So far: linear mixed-effects models

$\longrightarrow$

Before mixed models,

lets look at nonlinear regression

$\longrightarrow$

Future: nonlinear mixed-effects models

Roadmap

- Quick glance at the linear regression model
- Nonlinear regression model
- How to obtain the estimates: the Gauss-Newton algorithm
- Tests and confidence regions
- Examples with R

Linear Regression Models

Previously we have fitted, by least squares, the **linear regression model** which is of the type:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

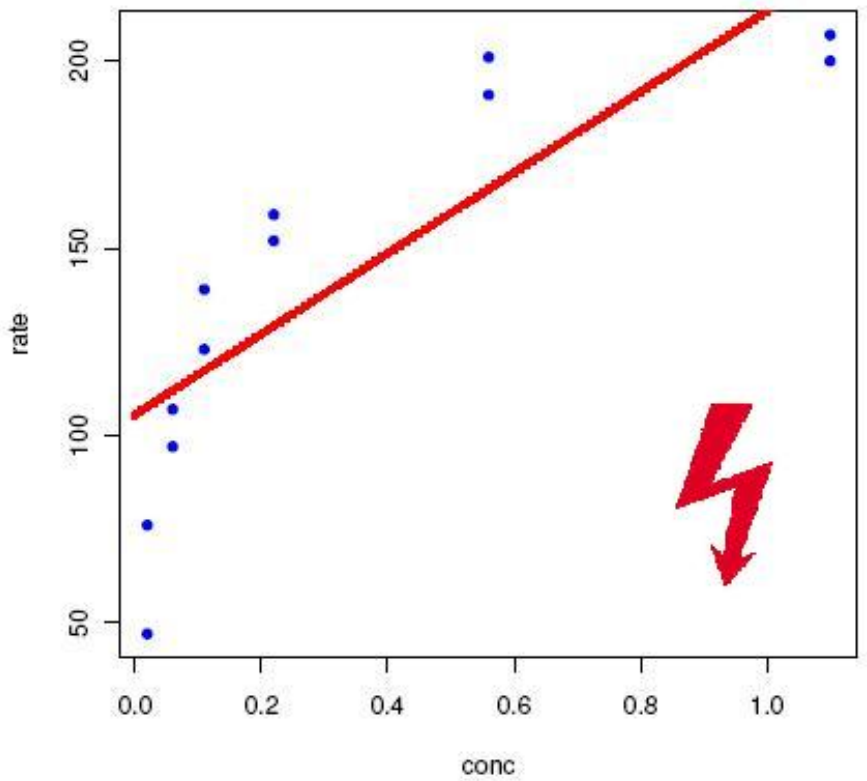
where

- $\mathbf{X}$ is the $n \times p$ matrix (p parameters and n observations) of explanatory variables, (or so-called **derivative matrix**);
- $\boldsymbol{\beta}$ is the vector of regression parameters to be estimated;
- $\boldsymbol{\epsilon}$ is a vector of random variables representing the disturbances, assumed to be iid normally distributed, $\mathbb{E}[\boldsymbol{\epsilon}] = 0$ and $Var(\boldsymbol{\epsilon}) = \sigma^2 \cdot \mathbf{I}$

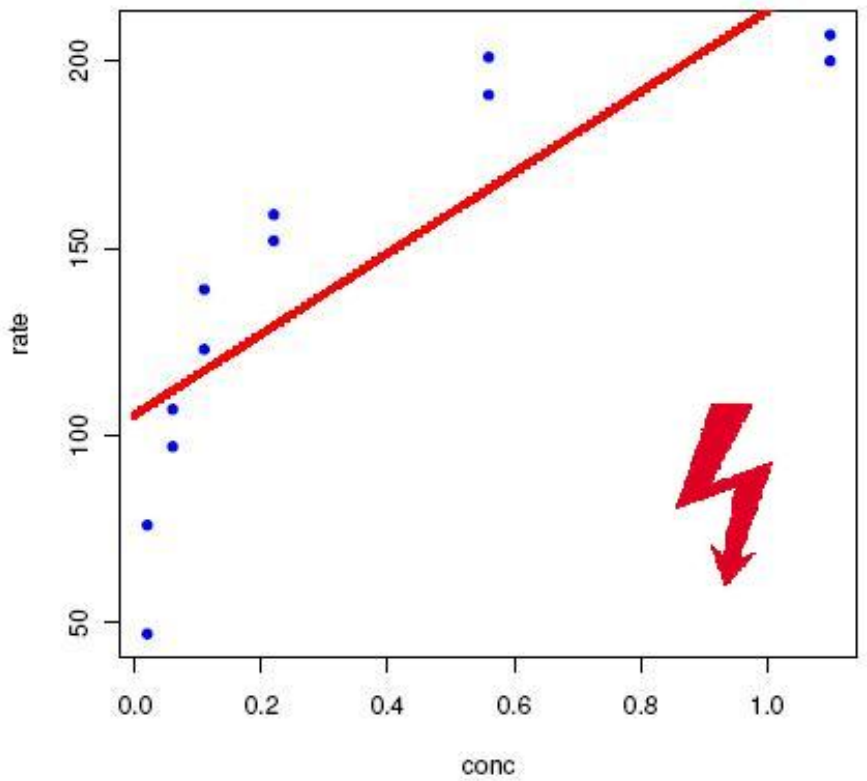
Model in vector notation:

$$Y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \epsilon_i \quad i = 1, \ldots, n$$

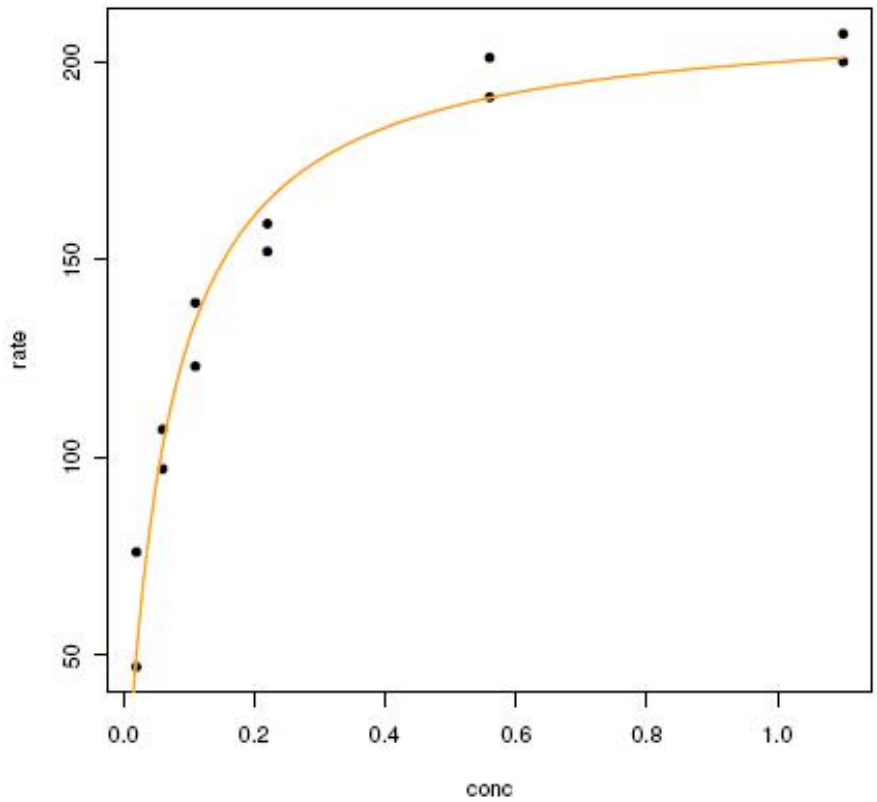
There are many situations in which a model of this form is not appropriate and too simple to represent the true relationship between the response variable Y and the predictor variables $X_1, \ldots X_p$.



There are many situations in which a model of this form is not appropriate and too simple to represent the true relationship between the response variable Y and the predictor variables $X_1, \ldots X_p$.



it would be better



First Definitions

The more general normal **nonlinear regression model** can be written

$$Y_i = f(\mathbf{x_i}, \theta) + \epsilon_\mathbf{i} \quad i = 1, \ldots, n$$

where f is the **expectation function** and $\mathbf{x_i}$ is a vector of associated regressor variables for the ith observation.
f is a nonlinear function of the parameter θ!

That is, for nonlinear models, at least one of the derivatives of the expectation function with respect to the parameters depends on at least one of the parameters.

Consider the vectors $\mathbf{x}_i$, $i = 1, 2, \ldots, n$ as fixed (observations) and concentrate on the dependance of the expected responses on $\boldsymbol{\theta}$. We create the n-dimensional vector $\boldsymbol{\eta}(\boldsymbol{\theta})$ with ith element

$$\eta_i(\boldsymbol{\theta}) = f(\mathbf{x}_i, \boldsymbol{\theta}) \qquad i = 1, 2, \ldots, n$$

and write the nonlinear regression model as

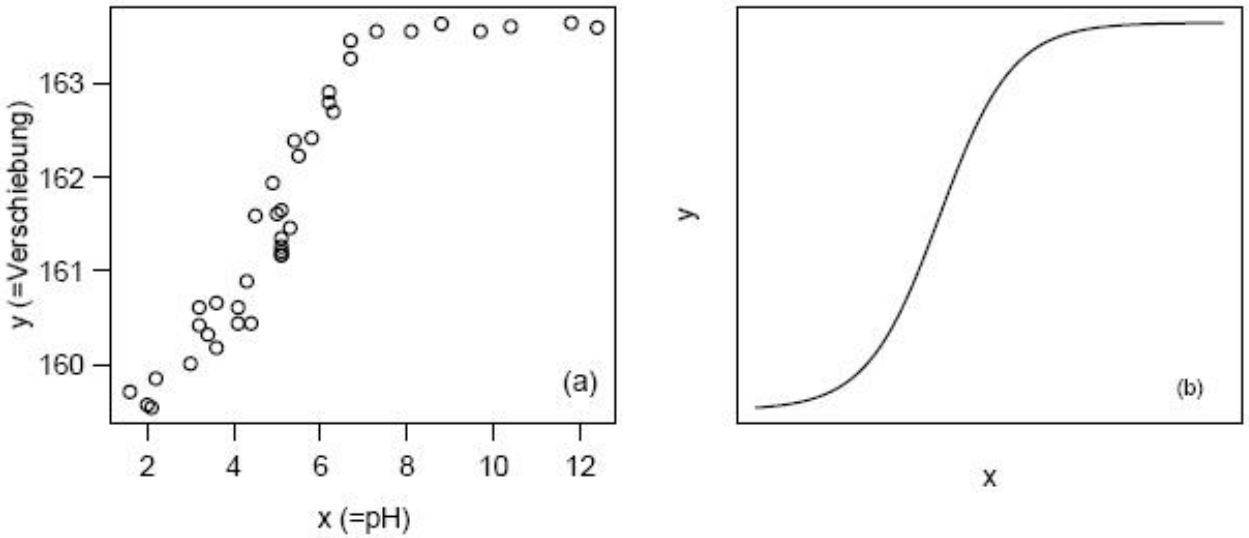
$$\mathbf{Y} = \boldsymbol{\eta}(\boldsymbol{\theta}) + \boldsymbol{\epsilon}$$

with $\boldsymbol{\epsilon}$ assumed to have a spherical normal distribution as in the linear model.

Examples for the nonlinear functions:

- From the membrane separation technology:

$$Y_i = f(\mathbf{x_i}, \boldsymbol{\theta}) + \boldsymbol{\epsilon_i} = \frac{\theta_1 + \theta_2 \cdot 10^{\theta_3 + \theta_4 x_i}}{1 + 10^{\theta_3 + \theta_4 x_i}} + \boldsymbol{\epsilon_i}$$



- Hill model (enzyme kinetics):

$$Y_i = \frac{\theta_1 + x_i^{\theta_3}}{(\theta_2 + x_i^{\theta_3})} + \boldsymbol{\epsilon_i}$$

For $\theta_3 = 1$ also known as Michaelis-Menton model.

- Mitscherlich function (growth analysis):

$$Y_i = \theta_1 + \theta_2 \cdot \exp(\theta_3 \cdot x_i) + \boldsymbol{\epsilon_i}$$

Transformably Linear Models

Many models can be transformed into a linear model. For example,

$$f(x, \boldsymbol{\theta}) = \theta_1 \; x^{\theta_2}$$

can be tranformed into a linear model:

$$\ln\Big(f(x, \boldsymbol{\theta})\Big) = \ln\theta_1 + \theta_2 \ln x = \beta_0 + \beta_1 \tilde{x},$$

where $\beta_0 = \ln\theta_1$ and $\tilde{x} = \ln x$.

We call such regression models transformably linear.

Attention!

Transformation of the data $\Rightarrow$ transformation of ϵ!

Initial model:

$$\ln Y_i = \beta_0 + \beta_1 \tilde{x} + \epsilon_i,$$

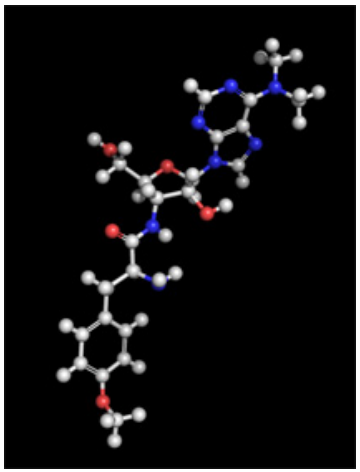
where ϵ_i additive, normally distributed. Transform it back, obtain:

$$Y_i = \theta_1 \cdot x^{\theta_2} \cdot \tilde{\epsilon}_i,$$

where $\tilde{\epsilon}_i$ perform as multiplicative and lognormally distributed!

Trasformation of the data is to enjoy with caution!

Example: Puromycin

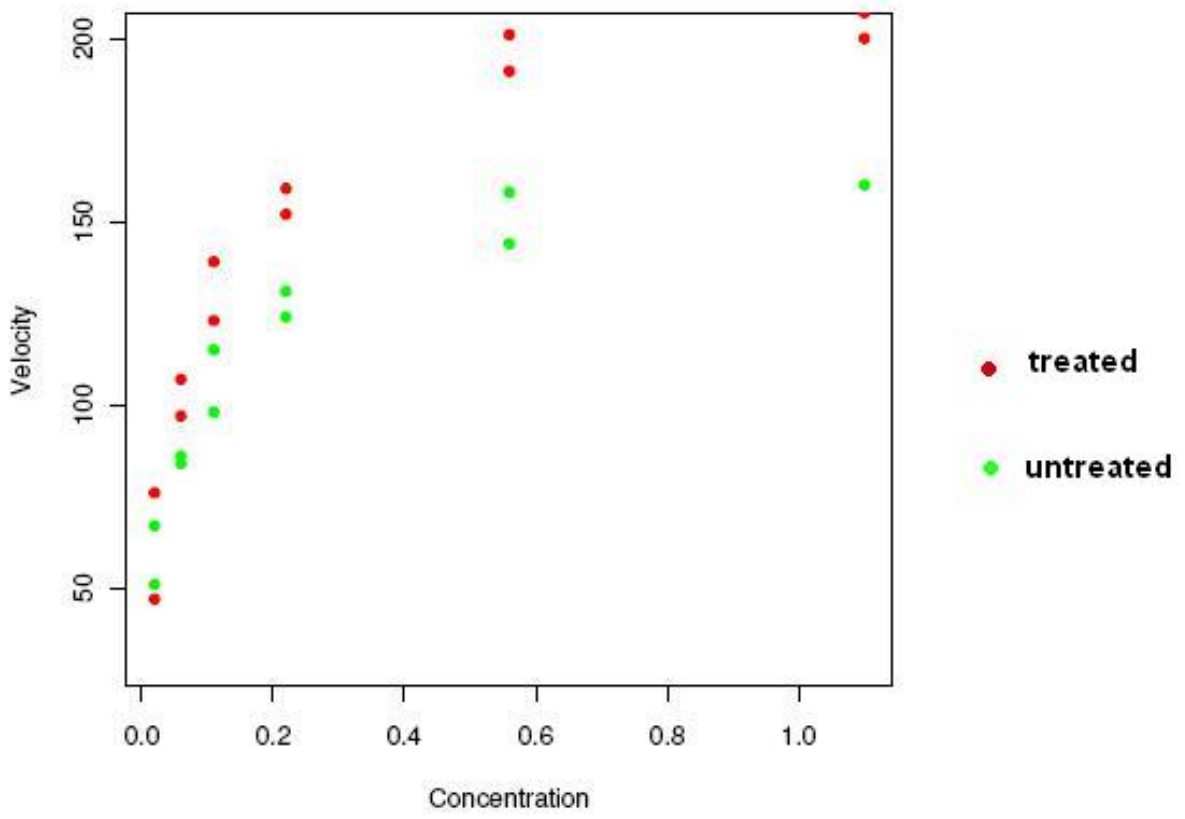


Description: Puromycin is an antibiotic used by scientists in bio-research to select cells modified by genetic engineering.

```
> Puromycin
   conc rate     state
1  0.02   76   treated
2  0.02   47   treated
3  0.06   97   treated
4  0.06  107   treated
5  0.11  123   treated
6  0.11  139   treated
7  0.22  159   treated
8  0.22  152   treated
9  0.56  191   treated
10 0.56  201   treated
11 1.10  207   treated
12 1.10  200   treated
13 0.02   67 untreated
14 0.02   51 untreated
15 0.06   84 untreated
16 0.06   86 untreated
17 0.11   98 untreated
18 0.11  115 untreated
19 0.22  131 untreated
20 0.22  124 untreated
21 0.56  144 untreated
22 0.56  158 untreated
23 1.10  160 untreated
```

Example: Puromycin

The data, for an enzyme treated and untreated:



Example: Puromycin

Mechanism of action: This is described by the Michaelis-Menten model for enzyme kinetics, which relates the initial velocity on an enzymatic reaction to the substrate concentration x trough the equation:

$$f(x, \boldsymbol{\theta}) = \frac{\theta_1 x}{\theta_2 + x}$$

Note, that for $x \to \infty$ we obtain the "asymptotic" reaction velocity equal to θ_1.

Puromycin

The Michaelis-Menten model can be transformed to a linear model:

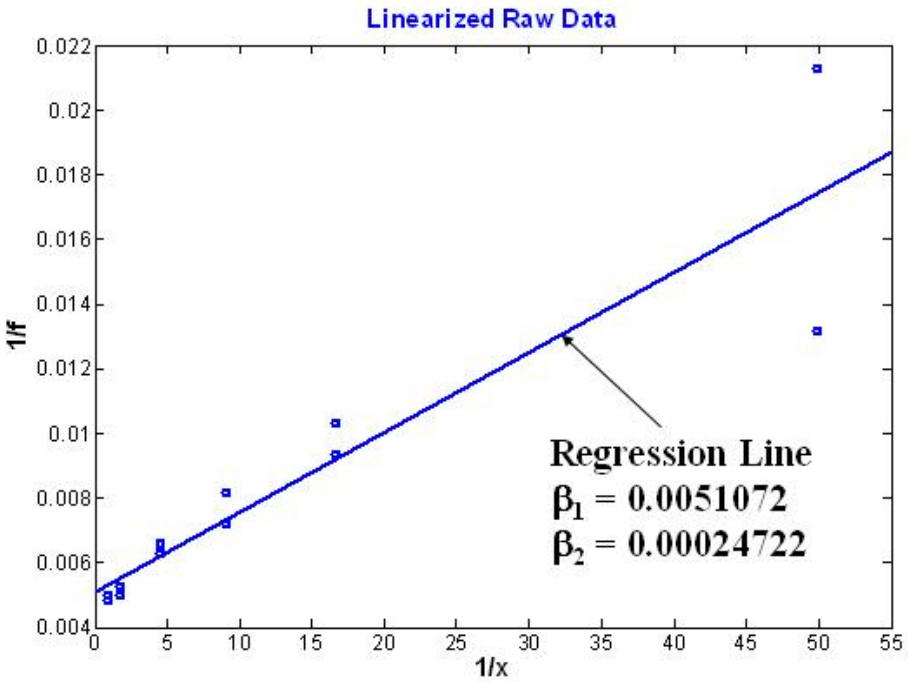
$$f(x, \boldsymbol{\theta}) = \frac{\theta_1 x}{\theta_2 + x}$$

$$\begin{aligned} \frac{1}{f} &= \frac{1}{\theta_1} + \frac{\theta_2}{\theta_1}\frac{1}{x} \\ &= \beta_1 + \beta_2 u \end{aligned}$$

$\leftarrow$ **linear model!**

Question: Is it a good option to do linear regression on this transformed data?

Puromycin



$$\frac{1}{f} = \beta_1 + \beta_2 \cdot \frac{1}{x}$$

The reciprocal data shows decidely nonconstant variance.

Transformation affected the disturbances!

Puromycin

Hence, we should use nonlinear regression on the original data.

- Can we bring the results obtained from the linear regression back into play?
- Yes, we can!

The Geometry of the Expectation Surface

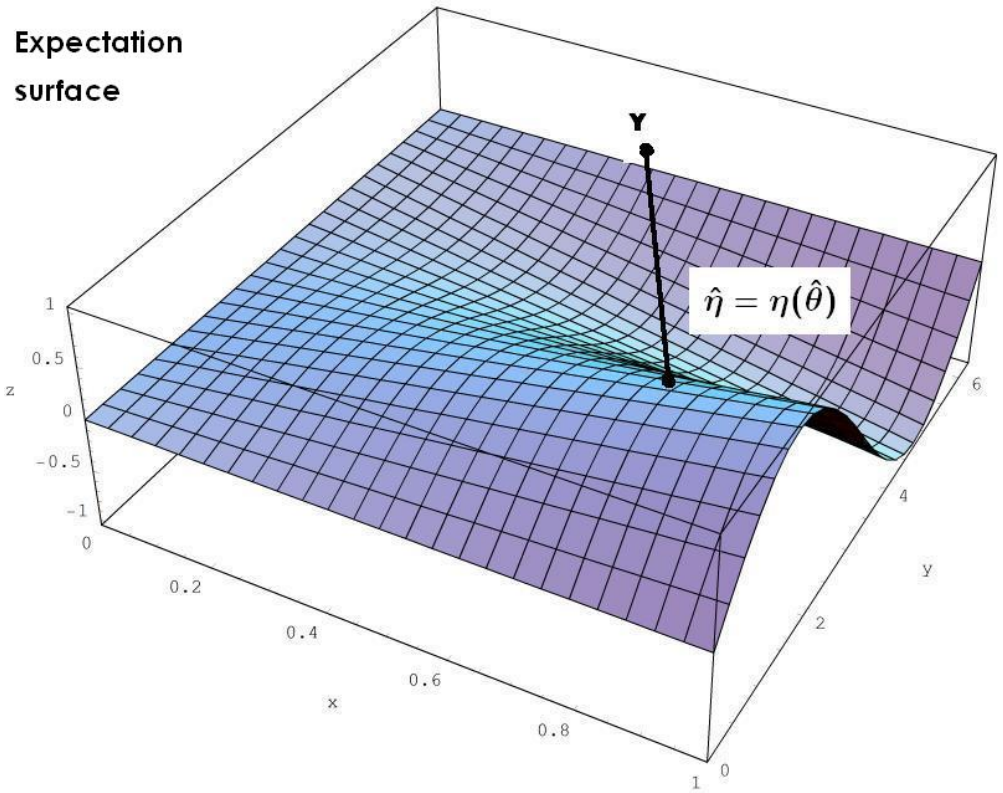
The n-dimensional vector $\boldsymbol{\eta}(\boldsymbol{\theta})$ defines a P-dimensional surface called the expectation surface in the N-dimensional response space. The least squares estimates correspond to the point on the expectation surface,

$$\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}(\hat{\boldsymbol{\theta}})$$

which is closest to $\mathbf{y} \in \mathbb{R}^n$. That is, $\hat{\boldsymbol{\theta}}$ minimizes the residual sum of squares

$$S(\boldsymbol{\theta}) = \|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\|^2$$

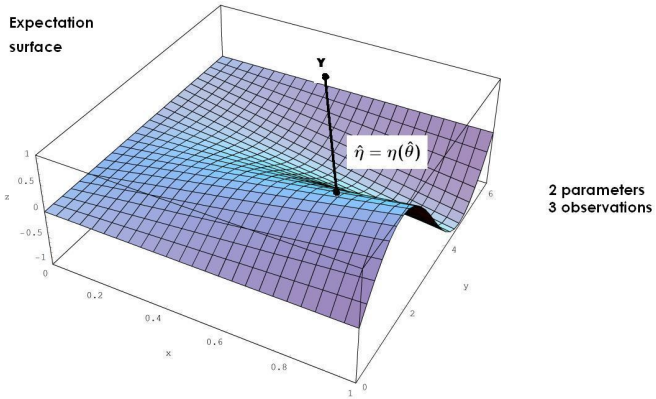
where $\boldsymbol{\eta}(\boldsymbol{\theta}) = f(\mathbf{x}, \boldsymbol{\theta})$.



2 parameters
3 observations

- find the point $\hat{\boldsymbol{\eta}}$ on the expectation surface which is closest to $\mathbf{y}$, and then
- determine the parameter vector $\hat{\boldsymbol{\theta}}$ which corresponds to the point $\hat{\boldsymbol{\eta}}$.

Determining the Least Squares Estimates



In the nonlinear case the two steps are very difficult, because:

- the expectation space is curved, often of finite extent (or, at least, has edges)
- we can map point easily only in one direction - from the parameter plane to the expectation surface. $=>$ It is extremely dificult to determine the parameter plane coordinates $\hat{\boldsymbol{\theta}}$ corresponding to that point $\hat{\boldsymbol{\eta}}$.

To overcome these difficulties, we use **iterative methods** to determine the least squares estimates $\hat{\boldsymbol{\theta}}$.

The Gauss-Newton Method: Idea

Goal: to iteratively improve an initial guess $\theta^{(0)}$ for θ and keep improving the estimates until there is no change.

- we expand the expectation function $f(\mathbf{x_i}, \boldsymbol{\theta})$ in a first order Taylor series about $\boldsymbol{\theta}^{(0)}$ as

$$f(\mathbf{x_i}, \boldsymbol{\theta}) \approx f(\mathbf{x_i}, \boldsymbol{\theta^{(0)}}) + A_{i1}(\theta_1 - \theta_1^{(0)}) + \cdots + A_{ip}(\theta_p - \theta_p^{(0)})$$

where

$$A_{ik} = \frac{\partial f(\mathbf{x_i}, \boldsymbol{\theta})}{\partial \theta_k} \Big|_{\boldsymbol{\theta}^{(0)}}, \quad k = 1, 2, \ldots, p$$

Incorporating all n cases, we write

$$\boldsymbol{\eta}(\boldsymbol{\theta}) \approx \boldsymbol{\eta}(\boldsymbol{\theta^{(0)}}) + \mathbf{A}^*(\boldsymbol{\theta} - \boldsymbol{\theta}^{(0)})$$

where $\mathbf{A}^*$ is the $n \times p$ derivative matrix with elements A_{ik}.

The Gauss-Newton Method

- adding the error terms:

$$\mathbf{Y} - \boldsymbol{\eta}(\boldsymbol{\theta}^{(0)}) = \mathbf{A}^* \cdot \underbrace{(\boldsymbol{\theta} - \boldsymbol{\theta}^{(0)})}_{:=\boldsymbol{\beta}} + \mathbf{E}$$

- on this stage we do linear regression and obtain $\boldsymbol{\beta}$.
- Put $\boldsymbol{\theta}^{(1)} := \boldsymbol{\theta}^{(0)} + \boldsymbol{\beta}$;
- The point $\boldsymbol{\eta}(\boldsymbol{\theta}^{(1)})$ should now be closer to $\mathbf{y}$ and so we move to this better parameter value $\boldsymbol{\theta}^{(1)}$ and perform another iteration.
- We iterate this process until there is no useful change in the elements of the parameter vector.

The Gauss-Newton Method

Short Overview: Gauss-Newton Method

1. Start with an initial value $\boldsymbol{\theta}^{(0)}$
2. Linear approximation of the expectation surface near $\boldsymbol{\eta}(\boldsymbol{\theta}^{(0)})$: matrix A
3. Local linear regression
4. Obtain better value $\boldsymbol{\theta}^{(1)}$
5. Iterate this process

Initial Values

Iterative methods need initial values to start from. How can we find them?

- From the previous knowledge of the experiment design
- Via transfomed linear model
- Analysis of the expectation function f

Puromycin

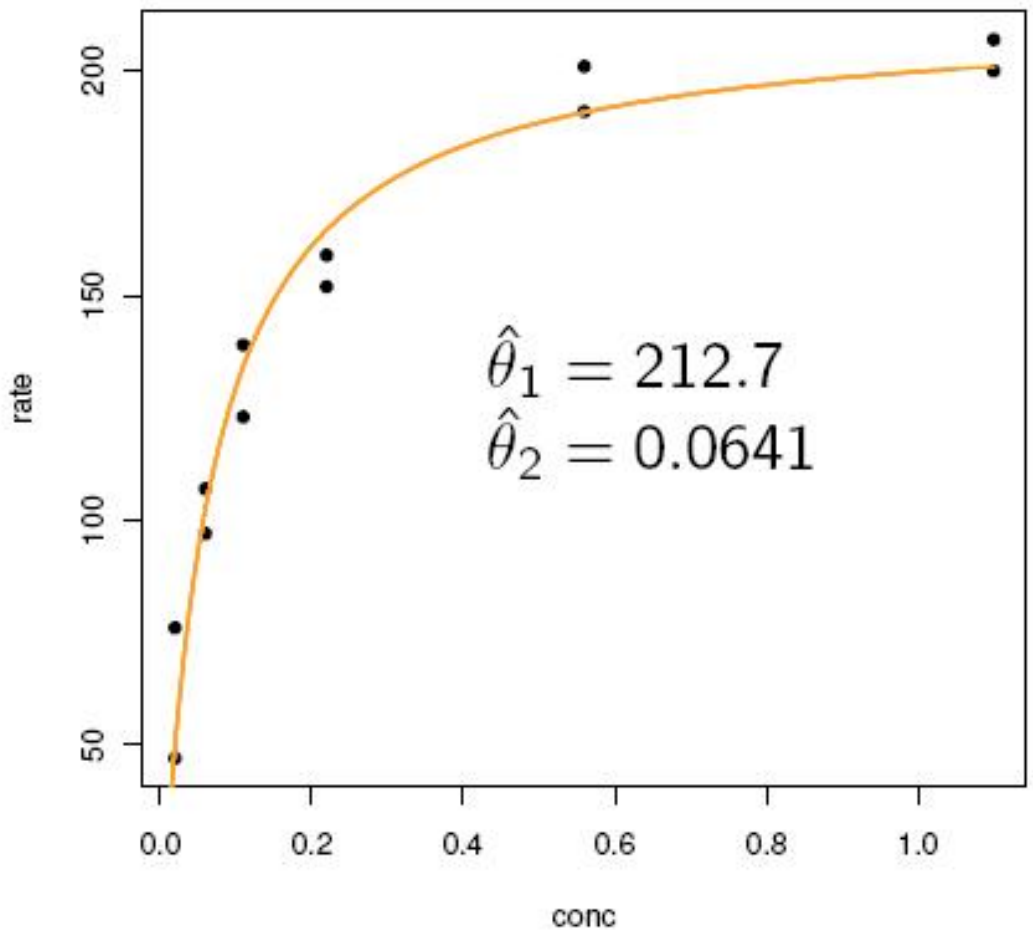
Now we are ready to do nonlinear regression on the Puromycin data. Take the initial values from the linear model:

$$\theta_1^{(0)} = 1/\hat{\beta}_0 = 196, \quad \theta_2^{(0)} = \hat{\beta}_1/\hat{\beta}_0 = 0.048$$

In R: use function **nls** to obtain nonlinear least squares estimates:

```
> r.nls<-nls(rate~(T1*conc)/(T2+conc),  start=list( T1=196, T2=0.048))
> summary(r.nls)

Formula: rate ~ (T1 * conc)/(T2 + conc)

Parameters:
    Estimate Std. Error t value Pr(>|t|)
T1 2.127e+02  6.947e+00  30.615 3.24e-11 ***
T2 6.412e-02  8.281e-03   7.743 1.57e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.93 on 10 degrees of freedom

Number of iterations to convergence: 6
Achieved convergence tolerance: 9.75e-07
```

Puromycin

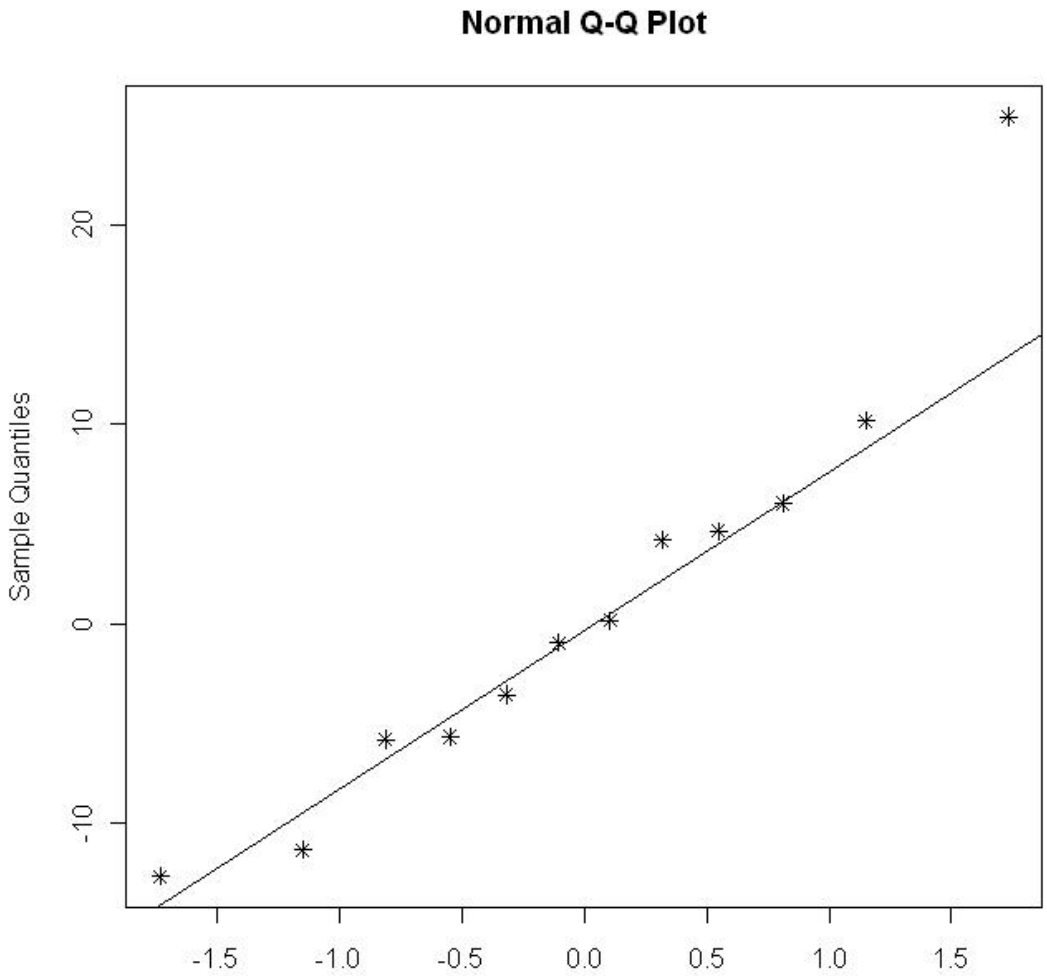


The fitted expectation function for the Puromycin data (treated)

$$f(x, \boldsymbol{\theta}) = \frac{\theta_1 x}{\theta_2 + x}$$

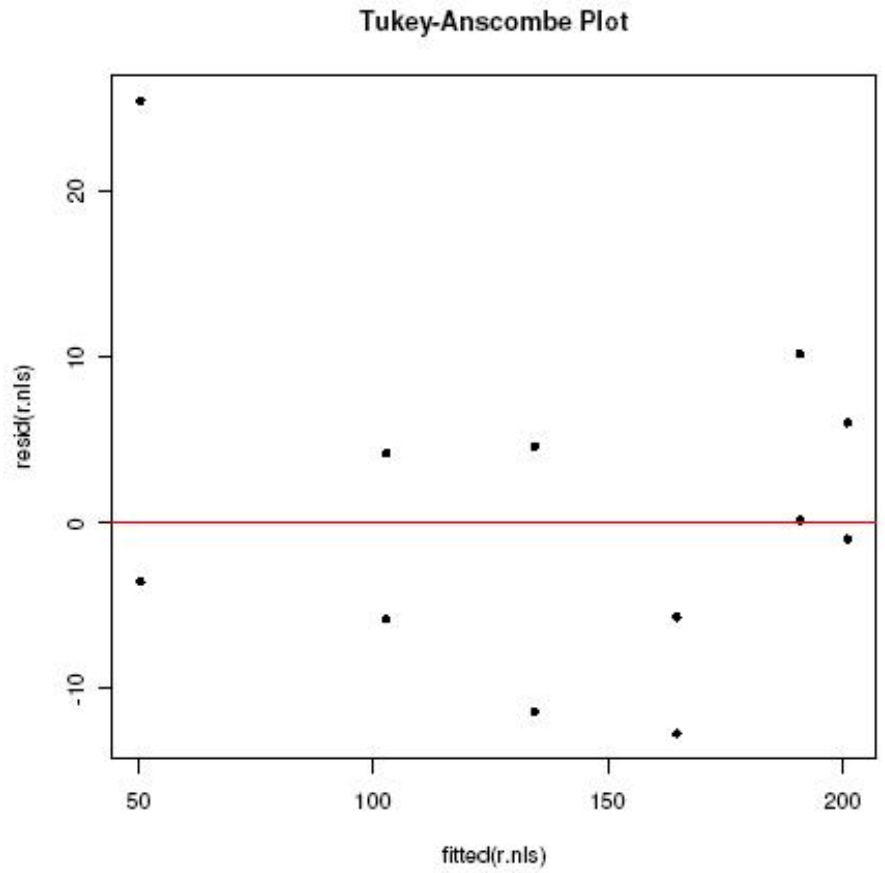
Puromycin

Model diagnostics: Normal QQ-Plot



Puromycin

Model diagnostics: Tukey-Anscombe-Plot



Approximate Tests and Confidence Regions

Goal: to construct confidence intervals for the parameters.

There are two approaches:

1. **Approximate test**, which is based on the linearization issue.

2. **Exact test**, where we compare the residual sums of squares $S(\theta^*)$ with $S(\theta)$.

Approximate Tests: Construction

Starting point: we have estimated $\hat{\boldsymbol{\theta}}$ with the LS method.

Question: is this estimate consistent with the observations?

Main assumption: The LS estimate $\hat{\theta}$ is asymptotically normal distributed:

$$\hat{\theta} \overset{a}{\sim} N\left(\theta, \frac{\mathbf{V}_\theta}{n}\right),$$

with asymptotic covariance matrix

$$\mathbf{V}_\theta = \sigma^2 \cdot \left(\mathbf{A}_\theta^T \cdot \mathbf{A}_\theta\right)^{-1}$$

Substitute $\mathbf{V}_\theta$ by its estimate:

$$\hat{\mathbf{V}}_\theta = \hat{\sigma}^2 \cdot \left(\mathbf{A}_{\hat{\theta}}^T \cdot \mathbf{A}_{\hat{\theta}}\right)^{-1} \text{ where } \hat{\sigma}^2 = \frac{1}{n-p} S(\hat{\theta})$$

Approximate Tests: Construction

Starting point: we have estimated $\hat{\boldsymbol{\theta}}$ with the LS method.

Question: is this estimate consistent with the observations?

Main assumption: The LS estimate $\hat{\boldsymbol{\theta}}$ is asymptotically normal distributed:

$$\hat{\boldsymbol{\theta}} \overset{a}{\sim} N\left(\boldsymbol{\theta}, \frac{\mathbf{V}_\theta}{n}\right),$$

with asymptotic covariance matrix

$$\mathbf{V}_\theta = \sigma^2 \cdot \left(\mathbf{A}_\theta^T \cdot \mathbf{A}_\theta\right)^{-1}$$

Substitute $\mathbf{V}_\theta$ by its estimate:

$$\hat{\mathbf{V}}_\theta = \hat{\sigma}^2 \cdot \left(\mathbf{A}_{\hat{\theta}}^T \cdot \mathbf{A}_{\hat{\theta}}\right)^{-1} \text{ where } \hat{\sigma}^2 = \frac{1}{n-p} S(\hat{\boldsymbol{\theta}})$$

$$H_{0,j}: \quad \theta_j = 0 \quad \textbf{against} \quad H_{A,j}: \quad \theta_j \neq 0$$

Nonlinear Model:

$$T_j = \frac{\hat{\theta}_j}{\sqrt{\hat{\sigma}^2 (A^T A)^{-1}_{jj}}} \overset{H_{0,j},\ as.}{\sim} t_{n-p}$$

Confidence region:

$$\hat{\theta}_j \pm \sqrt{\hat{\sigma}^2 (A^T A)^{-1}_{jj}} \cdot q^{t_{n-p}}_{1-\alpha/2},\ as.$$

Linear Model:

$$T_j = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 (X^T X)^{-1}_{jj}}} \overset{H_{0,j}}{\sim} t_{n-p}$$

Confidence region:

$$\hat{\beta}_j \pm \sqrt{\hat{\sigma}^2 (X^T X)^{-1}_{jj}} \cdot q^{t_{n-p}}_{1-\alpha/2}$$

Approximate Inference Bands for the Expected Response

The $1-\alpha$ approximate inference interval for the expected response is:

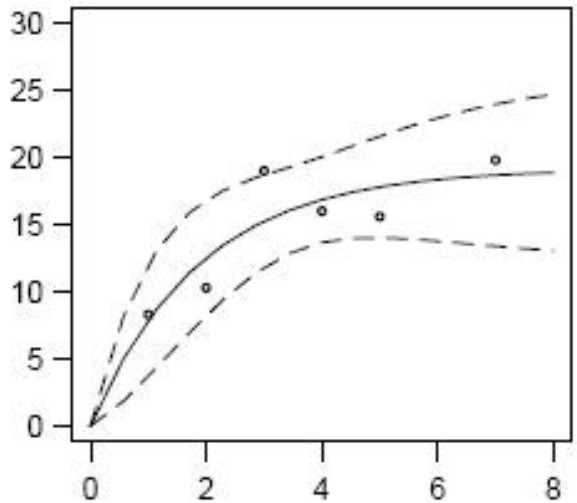
$$f(\mathbf{x}, \hat{\boldsymbol{\theta}}) \pm q_{1-\alpha/2}^{t_{n-p}} \cdot \quad \hat{\sigma}_{x_0}$$

where

$$\hat{\sigma}_{x_0} = \hat{\sigma} \sqrt{\hat{a}_0^T \left(A^T A\right)^{-1} \hat{a}_0}$$

with

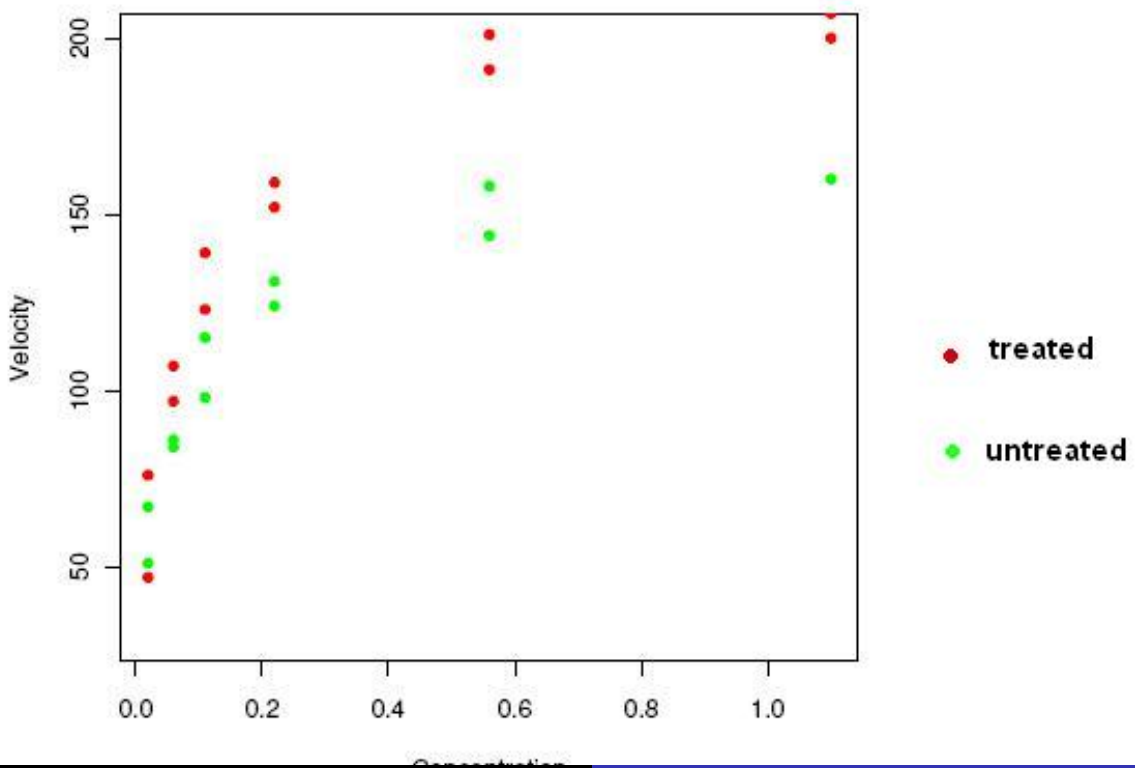
$$\hat{a}_0 = \frac{\partial f(\mathbf{x_0}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mid_{\hat{\boldsymbol{\theta}}}$$



Puromycin

In the Puromycin experiment, two blocks of experiments were run. In one the enzyme was treated with puromycin, and in the other the same enzyme was untreated.

Question: Does the Puromycin affect the parameters θ_1 or/and θ_2?



The combined model is then:

$$Y_i = \frac{(\theta_1 + \theta_3 z_i) \cdot x_i}{\theta_2 + \theta_4 \cdot z_i + x_i} + \epsilon_i$$

where z_i is the indicator variable:

$$z_i = \begin{cases} 0 & i\text{th case untreated, or} \\ 1 & \text{treated} \end{cases}$$

Parameters:

	Value	Std. Error	t value
θ_1	160.286	6.8964	23.24
θ_2	0.048	0.0083	5.76
θ_3	52.398	9.5513	5.49
θ_4	0.016	0.0114	1.44

Residual standard error: 10.4 on 19 degrees of freedom

We see, θ_4 could be zero, because its t-value 1.44 is smaller than the boundary $q^{t_{19}}_{0.975} = 2.09$.

The treatment has a significant effect on the velocity, which is expressed in θ_3:

The approximative 95% confidence interval for the parameter θ_3 is:

$$\hat{\theta}_3 \pm q^{t_{19}}_{0.975} \cdot \hat{\sigma} = 52.398 \pm 9.5513 \cdot 2.09 = [32.4, 72.4]$$

Exact Tests: Construction

The quality of the aprroximate test depends strongly on the quality of the linear approximation of the expectation surface. We can avoid this issue and construct tests which do not use linearisation components.

To test the null-hypothesis:
$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}^*$ or: $H_{0,j} : \theta_j = \theta_j^*$
we can compare the residual sums of squares $S(\boldsymbol{\theta}^*)$ with $S(\boldsymbol{\theta})$.

Exact Tests: F-Test

The test statistic is

$$T = \frac{n-p}{p} \cdot \frac{S(\boldsymbol{\theta}^*) - S(\hat{\boldsymbol{\theta}})}{S(\hat{\boldsymbol{\theta}})} \quad \overset{a}{\sim} \quad F_{p,n-p}.$$

This gives us the confidence region

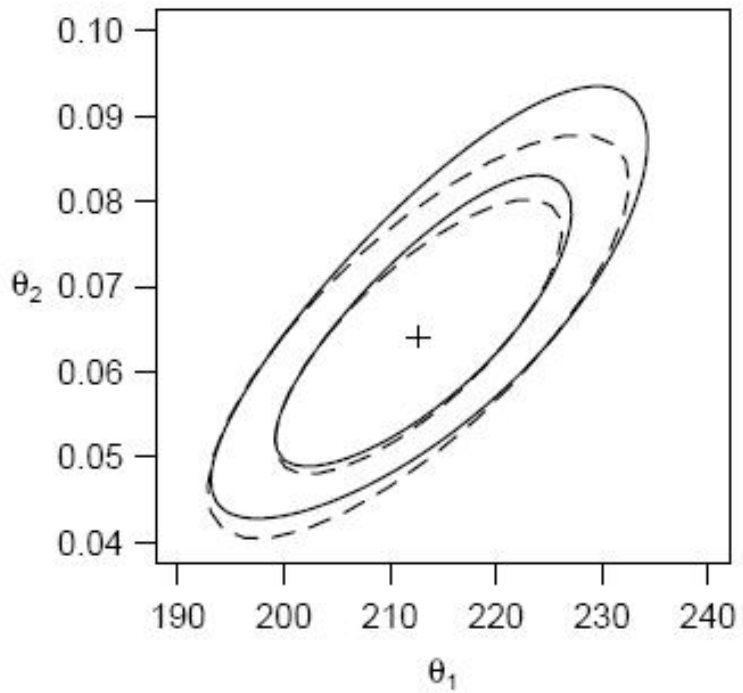
$$\left\{ \boldsymbol{\theta} \mid S(\boldsymbol{\theta}) \leq S(\hat{\boldsymbol{\theta}}) \left(1 + \frac{p}{n-p} \cdot q \right) \right\}$$

where $q = q_{1-\alpha}^{F_{p,n-p}}$ is the $(1-\alpha)$-quantile of the $F(p, n-p)$ distribution.

This test is **not** based on the issue of the linearization $=>$ it is more precise than the approximate test!

Exact Tests for p=2

When $P = 2$, we can determine a likelihood contour in $\boldsymbol{\theta}$ by standard contouring methods, that is, by evaluating $S(\boldsymbol{\theta})$ for a grid of $\boldsymbol{\theta}$ values and approximating the contour by straight line segments in the grid.



The figure shows nominal 80 and 95% likelihood contours for the Puromycin parameters. The dashed lines are the linear approximation ellipses, and the least square estimate is indicated by $+$.

Exact Test for a Single Parameter

For a single parameter:

$H_0 : \ \theta_j = \theta_j^*$

$$\tilde{T}_k = (n-p) \cdot \frac{\tilde{S}_k(\theta_k^*) - \tilde{S}(\hat{\theta})}{S(\hat{\theta})}$$

which is approx. $F_{1,n-p}$ - distributed.

We obtain the confidence interval by solving the equation $\tilde{T}_k = q_{1-\alpha/2}^{F_{1,n-p}}$ numerically for θ_k^*.

t-Test via F-Test

In the case of nonlinear regression the F-test is not equivalent to the f-test. But we can transform the the F-test into the t-test.

$$T_k(\theta_k^*) = \text{sign}\left(\hat{\theta}_k - \theta_k^*\right) \cdot (n-p)\frac{\sqrt{\tilde{S}_k(\theta_k^*) - S(\hat{\theta})}}{S(\hat{\theta})}$$

This statistic is approx. t_{n-p} - distributed.

Profile-t-Function

We introduce here some graphical techniques for assessing the severity of the nonlinearity in an estimation situation.

Look at the test statistic T_k as a function of θ_k, and call it **profile-t-function**:

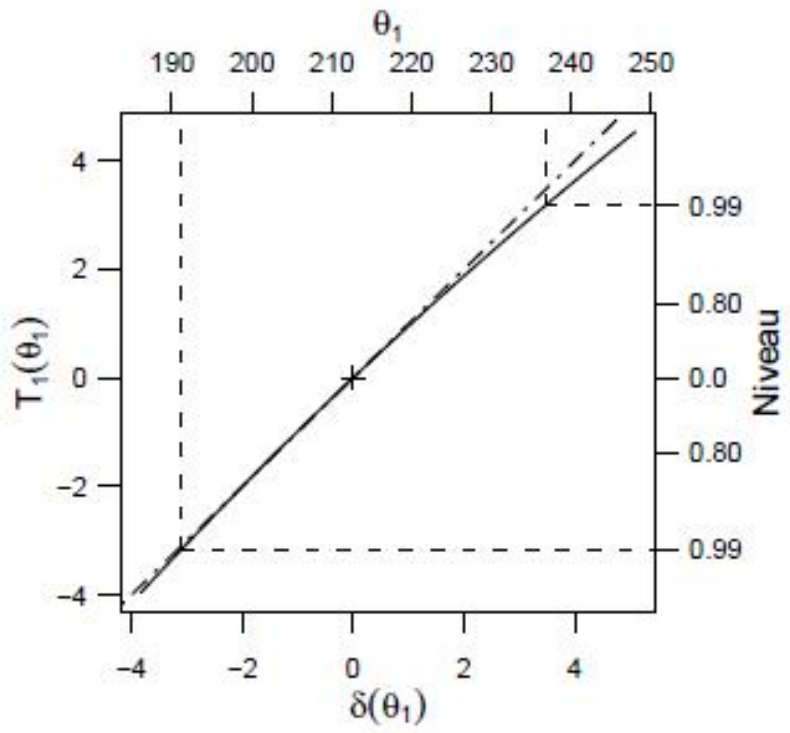
$$T_k(\theta_k) = \text{sign}\left(\hat{\theta}_k - \theta_k\right) \cdot (n-p) \cdot \frac{\sqrt{\tilde{S}_k(\theta_k) - S(\hat{\boldsymbol{\theta}})}}{S(\hat{\boldsymbol{\theta}})}$$

What do we obtain in the classical linear model?

What do we obtain in the nonlinear model?

Profile-t-Plot

The **profile-t-plot** provides exact likelihood intervals for individual parameters and, in addition, reveal how nonlinear the estimation sitiation is.



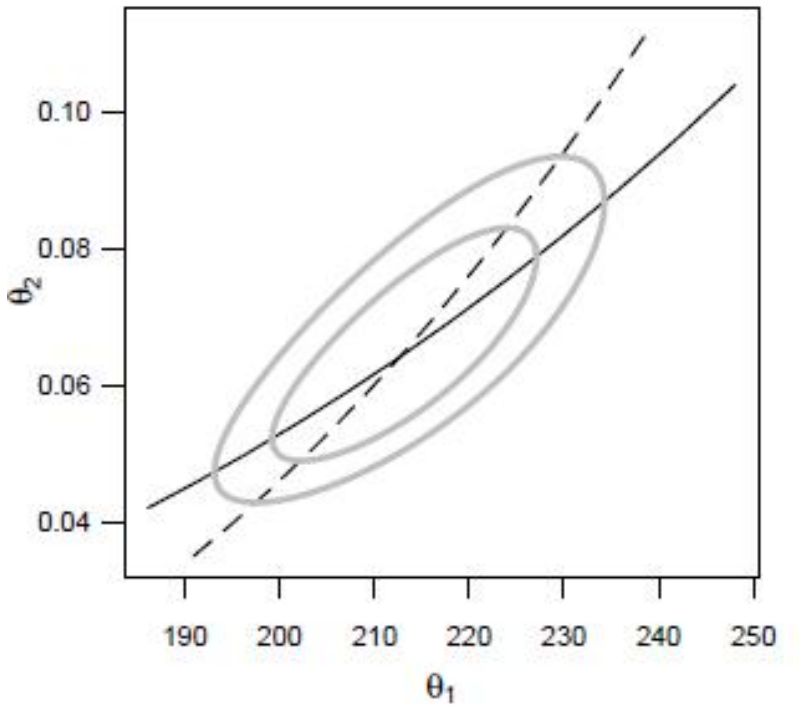
Usually we plot on the x-axis the standardised version of θ_k

$\delta_k(\theta_k) := \frac{\hat{\theta}_k - \theta_k^*}{se\langle(\hat{\theta}_k)\rangle}$ instead of θ_k.

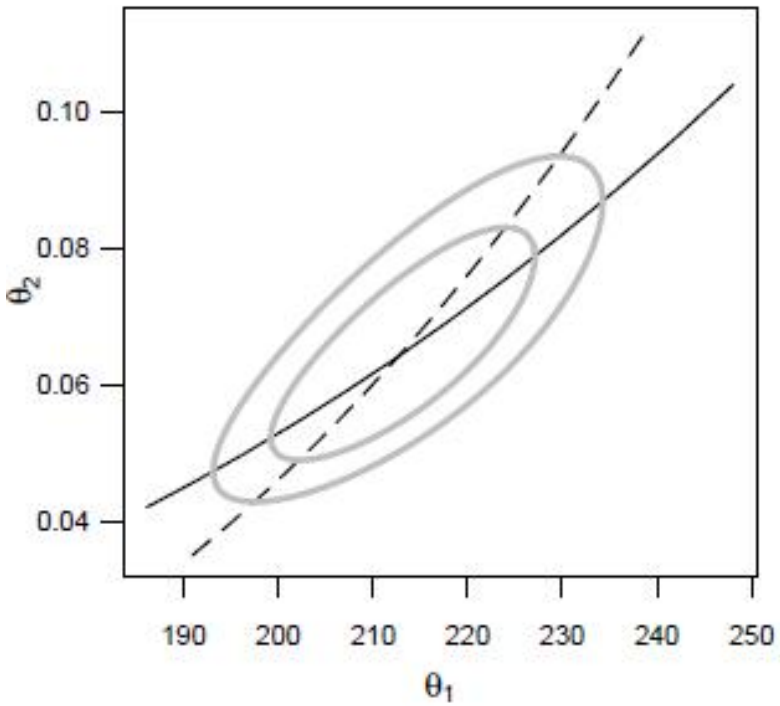
Interpretation?

Likelihood Profile Traces

Another tool is plotting the **likelihood profile traces**.
Here, we look at the estimated parameters $\hat{\theta}_j, j \neq k$, as a fuction $\hat{\theta}_j^{(k)}(\theta_k)$, while θ_k remains fixed. In addition, we may plot the likelihood contours, as in the following graphics: (Puromycin)



Interpretation?



traces are quite straight $\longrightarrow$ linear approximation is good.

Simular to the linear case: angle of intersection provides information about the correlation between the parameters.

Short Summary

- Look at the model: is it transformably linear?
- Collect the initial values
- Parameter estimates: the Gauss-Newton algorithm
- In R: function **nls**
- To construct inference regions: approximative or exact tests
- To look at the goodness of linearisation: **profile-t-plot**
- Correlation, how the estimate parameters interact: look at the **profile traces**

Good luck!

Any questions so far ??

Example with R: Chlorine

Chlorine decay from time $t = 0, ..., \infty$ in some product. Initial content of chlorine is 50%. The following nonlinear regression model is assumed:

$$\text{chlorine} = \alpha + (0.49 - \alpha) \cdot \exp(\beta \cdot \text{weeks} + \gamma) + \epsilon$$

Linearly transformable? First, we need $\alpha < \min(\text{chlorine})$. Why? Here, $\min(\text{chlorine}) = 0.38$, so take $\alpha = 0.37$ as initial value.

$$\text{chlorine} = \alpha + (0.49 - \alpha) \cdot \exp(\beta \cdot \text{weeks} + \gamma)$$

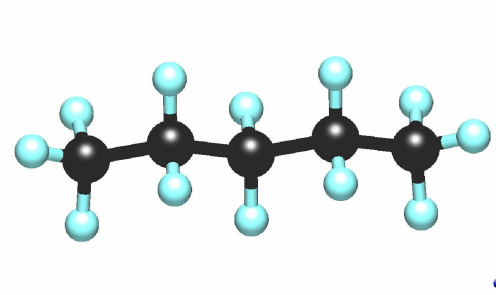
$$\frac{\text{chlorine} - 0.37}{0.49 - 0.37} = \exp(\beta \cdot \text{weeks} + \gamma)$$

$$\log\left(\frac{\text{chlorine} - 0.37}{0.12}\right) = \beta \cdot \text{weeks} + \gamma$$

That is, the transformed linear model is:

$$\log\left(\frac{\text{chlorine} - 0.37}{0.12}\right) = \beta \cdot \text{weeks} + \gamma + \epsilon^*$$

Example with R: n-Pentane



Data on the reaction rate of the catalytic isometrisation of *n*-pentane to isopentane versus the partial pressures of hydrogen, *n*-pentane and isopentane.

n-Pentane

The following nonlinear regression model (Hougen-Watson model) is assumed:

$$y = f(\mathbf{x}, \boldsymbol{\theta}) = \frac{\theta_1 \ \theta_3 \ (x_2 - x_3/1.632)}{1 + \theta_2 \ x_i + \theta_3 \ x_2 + \theta_4 \ x_3}$$

where

- y is the rate of reaction
- x_1 partial pressure of hydrogen
- x_2 partial pressure of n-pentane
- x_3 partial pressure of isopentane

Note: $\boldsymbol{\theta}$ has to be positive!

n-Pentane

This model is linear transformable:

$$y = \frac{\theta_1\ \theta_3\ (x_2 - x_3/1.632)}{1 + \theta_2\ x_1 + \theta_3 x_2 + \theta_4\ x_3}$$

$$\frac{1}{y} = \frac{1 + \theta_2 \cdot x_1 + \theta_3 \cdot x_2 + \theta_4 \cdot x_3}{\theta_1\theta_3 \cdot (x_2 - x_3/1.632)}$$

$$\frac{x_2 - x_3/1.632}{y} = \frac{1}{\theta_1\theta_3} + \frac{\theta_2}{\theta_1\theta_3} \cdot x_1 + \frac{1}{\theta_1} \cdot x_2 + \frac{\theta_4}{\theta_1\theta_3} \cdot x_3$$

$$\frac{x_2 - x_3/1.632}{y} = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3$$

n-Pentane

Denote

$$\tilde{y}_i = \frac{x_{i2} - x_{i3}/1.632}{y_i}$$

The linear model is then:

$$\tilde{y}_i = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \beta_3 \cdot x_{i3} + \epsilon_i$$

n-Pentane

Choose another parameter transformation:

$$y = \frac{\theta_1\ \theta_3\ (x_2 - x_3/1.632)}{1 + \theta_2\ x_1 + \theta_3 x_2 + \theta_4\ x_3}$$

Write

$$\phi_1 = \frac{1}{\theta_1\theta_3}, \qquad \phi_2 = \frac{\theta_2}{\theta_1\theta_3}, \qquad \phi_3 = \frac{1}{\theta_1}, \qquad \phi_4 = \frac{\theta_4}{\theta_1\theta_3}$$

Note, the ϕ_k corresponds to the β_k from above.

Then we obtain

$$y = \frac{x_2 - x_3/1.632}{\phi_1 + \phi_2 x_1 + \phi_3 x_2 + \phi_4 x_3}.$$

The End