Lecture 1: Course organization and introduction to linear models

Alan Lee

Department of Statistics
STATS 760 Lecture 1

March 9, 2017

Outline

Plan of today's lecture

In today's lecture we will discuss the course organization and then cover some introductory material on linear models.

My contact details

Office: Rm 367, building 303S

Phone: Extn 88749, DD 09-923-8749

Email: lee@stat.auckland.ac.nz

Course aims

- To give a survey of some statistical techniques you have not previously studied
- To develop skills in researching new techniques
- To give you same practice in speaking and writing about statistics

Course plan

- There will be very few lectures!!!
- Based around weekly group meetings
- Each of you will independently study 5 topics, with help from me
- In groups, you will write a short paper on one of the topics and give a verbal presentation on the topic

Possible topics

1. Linear Models
2. Data mining tools for regression and classification
3. Unsupervised learning
4. Survival analysis
5. Spatial statistics
6. Graphics
7. Other.....

Journals and meetings

- You will keep an individual journal in which you describe the things you have learnt, and give sample analyses.
- At our weekly meeting, we will discuss what you have done, and discuss what is to be done next.
- The final meeting will be an oral exam when I will ask you some questions to see what you have absorbed over the whole of your reading.

Assessment

Journal and meetings	20%
Paper and presentation	20%
Oral exam	20%
Take home test	20%
Assignments (5)	20%

How to approach a topic

For each topic, include in your journal

- An appreciation of why the technique is important and the kinds of applied problems that the technique can solve
- The types of data that call for the technique
- Details of the models being fitted
- The software that you need to implement the technique
- How to interpret the output from the computer runs
- Any diagnostic techniques that are important
- How the technique relates to others (e.g. how to glms relate to linear models, GAMS to glms)

Resources

Use the following resources to research your topics:

- Statistical Models in S
- Modern Applied Statistics with S-Plus (MASS)
- An Introduction to Statistical Learning (ISL)
- The Elements of Statistical Learning (ESL)
- R documentation
- Google
- other books and papers

Resources (cont)

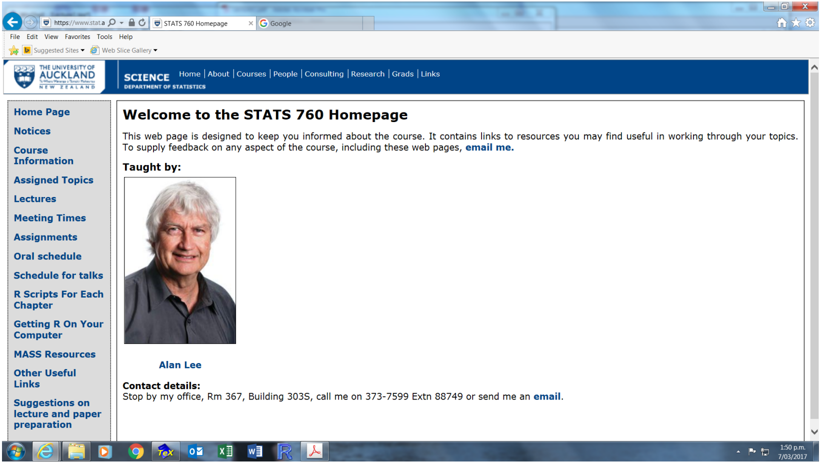
- Getting books out of the library can be frustrating if you all do it at once: Use e-books (MASS, ISL & ESL)
- Liaise with your classmates for sharing resources, Facebook page?, emails (I will post a list of class email addresses if the class agrees)
- There is an enormous amount of good stuff available over the net - Google madly!

Web page

`https://www.stat.auckland.ac.nz/ lee/760/`

- Course information
- Notice board - check frequently!
- Assigned topics
- R resources
- Suggestions on writing papers and giving presentations

Web page


https://www.stat.a
STATS 760 Homepage
Google
File Edit View Favorites Tools Help
Suggested Sites
Web Slice Gallery
THE UNIVERSITY OF AUCKLAND NEW ZEALAND
SCIENCE
DEPARTMENT OF STATISTICS
Home | About | Courses | People | Consulting | Research | Grads | Links
Home Page
Notices
Course Information
Assigned Topics
Lectures
Meeting Times
Assignments
Oral schedule
Schedule for talks
R Scripts For Each Chapter
Getting R On Your Computer
MASS Resources
Other Useful Links
Suggestions on lecture and paper preparation
Welcome to the STATS 760 Homepage
This web page is designed to keep you informed about the course. It contains links to resources you may find useful in working through your topics.
To supply feedback on any aspect of the course, including these web pages, email me.
Taught by:
Alan Lee
Contact details:
Stop by my office, Rm 367, Building 303S, call me on 373-7599 Extn 88749 or send me an email.
1:50 p.m.
7/03/2017

Housekeeping

Meetings: Once a week at times and places to be arranged, starting next week.

Presentations: In your group, decide on a topic by mid term break.

Topics: We can discuss at our first group meeting.

Class Rep

Any volunteers???

Linear models

- Review of 330
- The linear model: matrix formulation
- Least squares fitting: numerical details
- The geometric view
- Model matrices
- How R builds up the linear model matrix from the model formula
- Factors and parameterizations

Linear models resources

- Venables and Ripley Ch 6, particularly sections 6.2, 6.3
- Statistical Models in S, sections 2.2, 2.3, 2.4, 4.2
- STATS 330 lecture notes
  `https://www.stat.auckland.ac.nz/~stats330/`
- Reading list (000's of books on regression, linear models, see "Useful Links" on the web page for recommended titles)

R formulation

- Regression model: `y~x1+x2`
- Model with factors `y~A+B`
- Model with factors and continuous variables `y ~A*B*x1 + A*B*x2`

What do these mean? How do we interpret the output?

Linear models

These consist of two parts:

- A formula defining the mean of each observation, as a linear combination of explicit or implicit variables. Examples on previous slide.
- An assumption that the responses are independent, are normally distributed and have constant variance.

The particular linear combination being fitted is determined by the R formula, and some R system settings.

Numerical covariates

These are the easiest to understand. Suppose we have explanatory variables $x_1, \ldots, x_k$. Then the mean of an observation is assumed to be

$$\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

Our task is to estimate the β's. Geometrically, this amounts to fitting a plane through the data cloud. The formula in R is (say for k=2)

```
y~x1+x2.
```

We use least squares: the estimates of the β's are the values that minimise

$$\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_k)^2$$

Matrix formulation

Arrange data into a matrix and vector:

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \qquad X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix}$$

Then the mean of the observations is the vector $X\beta$ where the elements of β are $\beta_0, \ldots, \beta_k$. The estimates are obtained by minimizing

$$||y - X\beta||^2 = (y - X\beta)^T(y - X\beta)$$

Math stuff

- Vectors: $x^T = (x_1, x_2, \ldots, x_n)$, $y^T = (y_1, y_2, \ldots, y_n)$, (the ^{T} means we write the vectors as a row vector, rather than a column)
- Inner product: $(x, y) = x^T y = \sum_{i=1}^{n} x_i y_i$.
- Length: $||x|| = (x, x)^{1/2} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$.
- $||x + y||^2 = ||x||^2 + 2(x, y) + ||y||^2$.
- Orthogonality: x and y are orthogonal if $(x, y) = 0$.

Normal Equations

The minimizing β's ($\hat{\beta}$ say) satisfy

$$X^T X\hat{\beta} = X^T y$$

known as the normal equations.
Proof:

$$\begin{aligned}(y - X\beta)^T(y - X\beta) &= (y - X\hat{\beta} + X(\hat{\beta} - \beta))^T(y - X\hat{\beta} + X(\hat{\beta} - \beta)) \\ &= (y - X\hat{\beta})^T(y - X\hat{\beta}) \\ &\qquad + (X(\hat{\beta} - \beta))^T(X(\hat{\beta} - \beta) \\ &\geq (y - X\hat{\beta})^T(y - X\hat{\beta})\end{aligned}$$

Solving the equations

- We could calculate the matrix $X^T X$ and solve the set of linear equations. This is what SAS does - we can handle a large number of cases as $X^T X$ is only a $(k+1) \times (k+1)$ matrix. But this can be inaccurate: e.g. for polynomials.

- R uses the QR decomposition which is more accurate but requires all of X to be stored in the computer memory, unless some complicated programming is used.

QR decomposition

- Use the "QR decomposition" $X = QR$
- X is $n \times p$ and must have "full rank" (no column a linear combination of other columns)
- Q is $n \times p$ "orthogonal" (i.e. $Q^T Q =$ identity matrix)
- R is $p \times p$ "upper triangular" (all elements below the diagonal zero), all diagonal elements positive, so inverse exists
- Good computer algorithms exist for calculating this decomposition

Solving normal equations using QR

$$X^T X = R^T Q^T Q R = R^T R$$

$$X^T y = R^T Q^T y$$

Normal equations reduce to

$$R^T R\beta = R^T Q^T y$$

or

$$R\beta = Q^T y.$$

This is a triangular system and is easy to solve by *back-substitution*:

Back-substitution

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

Then $\beta_3 = c_3/r_{33}$,
$\beta_2 = (c_2 - r_{23}\beta_3)/r_{22}$,
$\beta_1 = (c_1 - r_{12}\beta_2 - r_{13}\beta_3)/r_{11}$.

What R does

When you run the R `lm` function, R figures out the matrix X and the vector y from the model formula, and then fits the model using the QR decomposition. Steps:

1. Form X and y.
2. Calculate the QR decomposition of X, calculate $c = Q^T y$.
3. Solve $R\beta = c$.
4. The solutions $\hat{\beta}$ are the numbers reported in the model summary.

Forming X

For a model formula say `y~x1+x2` where `x1` and `x2` are numeric variables, its a no-brainer:

1. Start with a column of 1's
2. Add columns corresponding to the explanatory variables.

It is more complicated for factors.

Factors

Suppose we have a model `y~A` where A is a factor with k levels. Typically this models the responses consisting of separate samples from k separate populations. The idea is that the mean just depends of the level of A, but is otherwise unspecified.

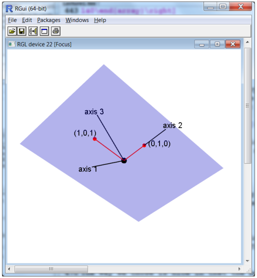
Example

Suppose we have a factor A with 2 levels, and three observations: y_1, y_2, y_3. The corresponding values of A are 1,2,1. This implies that the mean of the vector y is of the form $(\beta_1, \beta_2, \beta_1)$ since y_1 and y_3 have the same mean. We can write the mean of the observations in $X\beta$ form as

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \beta_1 + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \beta_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

Geometry

Another way of saying this is that the possible mean vectors are all possible linear combinations of the two vectors (1,0,1) and (0,1,0). Visually:



The set of all possible mean vectors is the plane determined by the two vectors. In math terms, it is the vector space spanned by the two vectors, which form a *basis* for the vector space.

Fitting in R

Now we have the X matrix we can go away and fit the model. However there is a problem:

We could describe the plane in an infinity of ways: any two non-parallel vectors lying in the plane will do. For each way, the coefficients will be different.

The way we chose is nice in that the betas have a natural interpretation: they are the means at factor levels 1 and 2.

We could have used (1,1,1) and (0,1,0) or (1,1,1) and (1,-1,1). These also have natural interpretations.

The point is there is no one way to do it.

Choosing a particular set of vectors that span the plane determines a particular *parameterization* for the means.

General case

Now suppose that the factor a has k levels, and there are an arbitrary number of observations. R has a built in mechanism for generating sets of vectors that span (are a basis for) the set of possible means, and thus lead to an X-matrix.

The general procedure is

1. Start with a column of 1's.
2. Add a dummy variable (k in all) for each level of the factor a (a dummy variable for level j has value one for all observations where $a = j$, and zero for the other observations)
3. We are not quite there, because the resulting X does not have linearly independent columns (the last k sum to the first), and the QR decomposition won't work.
4. Replace the last k columns (thought of as a matrix X_a) by $X_a C_a$ where C_a is a $(k \times (k-1)$ "contrast matrix" whose columns are linearly independent and linearly independent of the column of ones.
5. R has some built-in contrast matrices. We shall look at the "treatment" contrasts (mainly used in 330) and the "sum" contrasts (mainly used in 20x).

Treatment contrasts

Take k=3. Then the C-matrix is

$$C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

In general, the effect of applying the C matrix is to delete the first dummy variable, leaving the others unchanged. To see the general form of the C-matrix, start R and type (e.g. for k=4)

```
> k=4
> contr.treatment(k)
  2 3 4
1 0 0 0
2 1 0 0
3 0 1 0
4 0 0 1
```

Applying treatment contrasts

$$
\begin{array}{c c c c}
\multicolumn{4}{c}{[1 : X_a]} \\
1 & 1 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
1 & 1 & 0 & 0 \\
\hline
1 & 0 & 1 & 0 \\
\vdots & \vdots & \vdots & \\
1 & 0 & 1 & 0 \\
\hline
1 & 0 & 0 & 1 \\
\vdots & \vdots & \vdots & \\
1 & 0 & 0 & 1
\end{array}
\quad
\begin{array}{c}
\\ \\ \longrightarrow \\ \\ \\ \longrightarrow \\ \\ \\ \longrightarrow \\ \\
\end{array}
\quad
\begin{array}{c c c}
\multicolumn{3}{c}{[1 : X_a C_a]} \\
1 & 0 & 0 \\
\vdots & \vdots & \vdots \\
1 & 0 & 0 \\
\hline
1 & 1 & 0 \\
\vdots & \vdots & \vdots \\
1 & 1 & 0 \\
\hline
1 & 0 & 1 \\
\vdots & \vdots & \vdots \\
1 & 0 & 1
\end{array}
\begin{array}{l}
\\
\left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{observations at level 1} \\
\left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{observations at level 2} \\
\left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{observations at level 3}
\end{array}
$$

Treatment contrasts: interpretation

If $\beta_1, \beta_2, \beta_3$ are the coefficients, then

- Mean response at level 1 is β_1
- Mean response at level 2 is $\beta_1 + \beta_2$
- Mean response at level 3 is $\beta_1 + \beta_3$

Thus,

- β_1 is interpreted as the baseline (level 1) mean
- β_2 is interpreted as the offset for level 2 (difference between levels 1 and 2)
- β_3 is interpreted as the offset for level 3 (difference between levels 1 and 3)

Choosing contrasts: Treatment contrasts

These are the default in R, if this has been changed type

```
> options(contrasts=c("contr.treatment", "contr.poly"))
> options("contrasts")
$contrasts
[1] "contr.treatment" "contr.poly" }
```

to reset them.

Sum contrasts

Take k=3. Then the C-matrix is

$$C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix}$$

In general, the effect of applying the C matrix is to delete the last dummy variable and change the elements in rows corresponding to the last level of the factor to -1. To see the general form of the C-matrix in this case, start R and type (e.g. for $k = 4$)

```
> k=4
> contr.sum(k)
  [,1] [,2] [,3]
1    1    0    0
2    0    1    0
3    0    0    1
4   -1   -1   -1
```

Applying sum contrasts

$$[1 : X_a] \qquad\qquad [1 : X_a C_a]$$

$$\begin{array}{cccc} 1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \\ 1 & 1 & 0 & 0 \\ \hline 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \\ 1 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \\ 1 & 0 & 0 & 1 \end{array} \quad \begin{array}{c} \\ \longrightarrow \\ \\ \\ \longrightarrow \\ \\ \\ \longrightarrow \\ \\ \end{array} \quad \begin{array}{ccc} 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \\ \hline 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \\ \hline 1 & -1 & -1 \\ \vdots & \vdots & \vdots \\ 1 & -1 & -1 \end{array} \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{observations at level 1} \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{observations at level 2} \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{observations at level 3} \end{array}$$

Sum contrasts: interpretation

If $\beta_1, \beta_2, \beta_3$ are the coefficients, then

- Mean response at level 1 is $\beta_1 + \beta_2$
- Mean response at level 2 is $\beta_1 + \beta_3$
- Mean response at level 3 is $\beta_1 - \beta_2 - \beta_3$

Thus,

- β_1 is interpreted as the overall mean of the 3 means
- β_2 is interpreted as the offset for level 1 (difference between overall mean and level 1)
- β_3 is interpreted as the offset for level 2 (difference between overall mean and level 2)

Choosing contrasts: Sum contrasts

To select sum contrasts, type

```
> options(contrasts=c("contr.sum", "contr.poly"))
> options("contrasts")
$contrasts
[1] "contr.sum" "contr.poly" }
```

Interpreting parameters

To interpret the parameters correctly, we need to know the relationship between the means of the observations and the parameters. We can use R to help us understand this. Steps:

1. Delete the duplicate rows from the X-matrix. (observations corresponding to duplicate rows have the same mean). Call the result X_*.
2. Let μ be the vector of means in the same order as the rows of X_*. Then the relationship between μ and β is $\mu = X_*\beta$. Inverting this, we can express β in terms of μ as $\beta = (X_*^T X_*)^{-1} X_*^T \mu$. We can look at the matrices X_* and $(X_*^T X_*)^{-1} X_*^T$ to help us understand the relationships.

Example: Single factor, see 330 Lecture 18

In an experiment to study the effect of carcinogenic substances, six different substances were applied to cell cultures. The response variable (ratio) is the ratio of damaged to undamaged cells, and the explanatory variable (treatment) is the substance. Data (in R data frame `carcin.df`) are

ratio	treatment
0.08	control
+ 49 other control obs	
0.08	choralhydrate
+ 49 other choralhydrate obs	
0.10	diazapan
+ 49 other diazapan obs	
0.10	hydroquinone
+ 49 other hydroquinine obs	
0.07	econidazole
+ 49 other econidazole obs	
0.17	colchicine

R code

```
> carcin.df = read.table(file.choose(), header=TRUE)
> means = c("control","chloralhydrate","diazapan",
    "hydroquinone", "econidazole", "colchicine")
> carcin.df$treatment = factor(carcin.df$treatment,
     levels=means)
> cancer.lm=lm(ratio~treatment, data=carcin.df)
> X = model.matrix(cancer.lm)
> Xstar = X[c(1,51,101,151,201,251),]
> dimnames(Xstar)[[1]] = means
```

X_*

```
> Xstar
              (Intercept) treatmentchloralhydrate treatmentdiazapan treatmenthydroquinone
control                 1                       0                 0                     0
chloralhydrate          1                       1                 0                     0
diazapan                1                       0                 1                     0
hydroquinone            1                       0                 0                     1
econidazole             1                       0                 0                     0
colchicine              1                       0                 0                     0
             treatmenteconidazole treatmentcolchicine
control                         0                   0
chloralhydrate                  0                   0
diazapan                        0                   0
hydroquinone                    0                   0
econidazole                     1                   0
```

$(X_*^T X_*)^{-1} X_*^T$

```
> round(solve(t(Xstar)%*%Xstar)%*%t(Xstar))
                          control chloralhydrate diazapan hydroquinone econidazole colchicine
(Intercept)                     1              0        0            0           0          0
treatmentchloralhydrate        -1              1        0            0           0          0
treatmentdiazapan              -1              0        1            0           0          0
treatmenthydroquinone          -1              0        0            1           0          0
treatmenteconidazole           -1              0        0            0           1          0
treatmentcolchicine            -1              0        0            0           0          1
```

Two factors

For the model `y ~ a + b` where `a` and `b` are factors, the X-matrix is formed as follows

- Start with column of 1's
- Add $X_a C_a$
- Add $X_b C_b$

For the model `y ~a * b` where `a` and `b` are factors, the X-matrix is formed as above but further columns are added: Every column of $X_a C_a$ is multiplied elementwise with every column of $X_b C_b$.

Example: Experiment to study weight gain in rats

Response is weight gain over a fixed time period This is modelled as a function of diet (Beef, Cereal, Pork) and amount of feed (High, Low) See 330 Lecture 18. The data are

```
> diets.df
   gain source level
1    73   Beef  High
2    98 Cereal  High
3    94   Pork  High
4    90   Beef   Low
5   107 Cereal   Low
6    49   Pork   Low

7   102   Beef  High
8    74 Cereal  High
9    79   Pork  High
```

Example (continued)

There are 6 distinct means:

	Source		
Level	Beef	Cereal	Pork
High	μ_1	μ_2	μ_3
Low	μ_4	μ_5	μ_6

Example (continued)

```
> rats.lm<-lm(gain~source + level, data=diets.df)
> X<-model.matrix(rats.lm)[1:6,]
> X
  (Intercept) sourceCereal sourcePork levelLow
1           1            0          0        0
2           1            1          0        0
3           1            0          1        0
4           1            0          0        1
5           1            1          0        1
6           1            0          1        1
```

Example (continued)

```
> rats.lm<-lm(gain~source + level, data=diets.df)
> X<-model.matrix(rats.lm)[1:6,]
> coef.mat<-solve(t(X)%*%X)%*%t(X)
> round(6*coef.mat)
              1  2  3  4  5  6
(Intercept)   4  1  1  2 -1 -1
sourceCereal -3  3  0 -3  3  0
sourcePork   -3  0  3 -3  0  3
levelLow     -2 -2 -2  2  2  2
```

Example (continued)

```
> rats.lm<-lm(gain~source * level, data=diets.df)
> X<-model.matrix(rats.lm)[1:6,]
  (Intercept) sourceCereal sourcePork levelLow sourceCereal:levelLow sourcePork:levelLow
1           1            0          0        0                     0                   0
2           1            1          0        0                     0                   0
3           1            0          1        0                     0                   0
4           1            0          0        1                     0                   0
5           1            1          0        1                     1                   0
6           1            0          1        1                     0                   1
```

Example (continued)

```
> rats.lm<-lm(gain~source * level, data=diets.df)
> X<-model.matrix(rats.lm)[1:6,]
> coef.mat<-solve(t(X)%*%X)%*%t(X)
> round(coef.mat)
                      1  2  3  4 5 6
(Intercept)           1  0  0  0 0 0
sourceCereal         -1  1  0  0 0 0
sourcePork           -1  0  1  0 0 0
levelLow             -1  0  0  1 0 0
sourceCereal:levelLow 1 -1  0 -1 1 0
sourcePork:levelLow   1  0 -1 -1 0 1
```

One continuous, one factor

Lathe example (330 Lecture 17)

Consider an experiment to measure the rate of metal removal in a machining process on a lathe. The rate depends on the speed setting of the lathe (fast, medium or slow, a categorical measurement) and the hardness of the material being machined (a continuous measurement). Data are

```
   hardness setting rate
1       120    slow   68
2       140    slow   90
3       150    slow   98
4       125    slow   77
5       136    slow   88
6       165  medium  122
7       140  medium  104
8       120  medium   75
9       125  medium   84
10      133  medium   95
11      175    fast  138
12      132    fast  102
13      124    fast   93
14      141    fast  112
15      130    fast  100
```

Two models

`rate ~setting + hardness`

The X-matrix is formed as follows

- Start with column of 1's
- Add $X_{Setting} C_{Setting}$
- Add single column for hardness

`rate ~setting * hardness`

The X-matrix is formed as follows

- Start with column of 1's
- Add $X_{Setting} C_{Setting}$
- Add single column for hardness
- Add further columns by multiplying each column of $X_{Setting} C_{Setting}$ elementwise by hardness

Parallel lines model

```
> model1 = lm(rate~setting + hardness, data=metal.df)
> model.matrix(model1)
   (Intercept) settingmedium settingslow hardness
1            1             0           1      120
2            1             0           1      140
3            1             0           1      150
4            1             0           1      125
5            1             0           1      136
6            1             1           0      165
7            1             1           0      140
8            1             1           0      120
9            1             1           0      125
10           1             1           0      133
11           1             0           0      175
12           1             0           0      132
13           1             0           0      124
```

Non-Parallel lines model

```
> model2 = lm(rate~setting * hardness, data=metal.df)
> model.matrix(model2)
   (Intercept) settingmedium settingslow hardness settingmedium:hardness settingslow:hardness
1            1             0           1      120                      0                 120
2            1             0           1      140                      0                 140
3            1             0           1      150                      0                 150
4            1             0           1      125                      0                 125
5            1             0           1      136                      0                 136
6            1             1           0      165                    165                   0
7            1             1           0      140                    140                   0
8            1             1           0      120                    120                   0
9            1             1           0      125                    125                   0
10           1             1           0      133                    133                   0
11           1             0           0      175                      0                   0
12           1             0           0      132                      0                   0
13           1             0           0      124                      0                   0
14           1             0           0      141                      0                   0
15           1             0           0      130                      0                   0
```