Chapter 1

Linear Models with Fixed Effects Only

The basic linear model with fixed effects only is:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$
$$\boldsymbol{\varepsilon} \sim \mathrm{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$$

$\boldsymbol{y}, \boldsymbol{\varepsilon} \in \mathbb{R}^n$, $\boldsymbol{\beta} \in \mathbb{R}^p$, $\boldsymbol{X} \in \mathbb{R}^{n \times p}$ with rank p.

Example 1.1. Simple linear regression model.

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \qquad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

Single-observation representation:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad \varepsilon_i \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, \sigma^2) \qquad \triangle$$

Example 1.2. One-way ANOVA model. Cell means form:

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\ \vdots \\ 1\end{matrix}}\right\} n_1 \text{ rows} \\ \left.\vphantom{\begin{matrix}0\\ \vdots \\ 0\end{matrix}}\right\} n_2 \text{ rows} \\ \\ \left.\vphantom{\begin{matrix}0\\ \vdots \\ 0\end{matrix}}\right\} n_M \text{ rows} \end{matrix} \qquad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_M \end{bmatrix}$$

Single-observation representation using a double index:

$$y_{ij} = \beta_i + \varepsilon_{ij} \qquad \varepsilon_{ij} \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$$

Alternative parametrization, using the first group as the reference (as in R's "treatment contrasts"):

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \\ 1 & 0 & 0 & & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & & 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\\vdots\\1\end{matrix}}\right\} n_1 \text{ rows} \\ \left.\vphantom{\begin{matrix}1\\\vdots\\1\end{matrix}}\right\} n_2 \text{ rows} \\ \left.\vphantom{\begin{matrix}1\\\vdots\\1\end{matrix}}\right\} n_3 \text{ rows} \\ \\ \left.\vphantom{\begin{matrix}1\\\vdots\\1\end{matrix}}\right\} n_M \text{ rows} \end{matrix} \qquad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_M \end{bmatrix}$$

Single-observation representation using a double index:

$$y_{ij} = \beta_0 + \beta_i + \varepsilon_{ij} \qquad \varepsilon_{ij} \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$$

with $\beta_1 = 0$.   △

Example 1.3. Rail data (P&B p. 4ff) – see Figure 1.1: Travel times of ultrasonic waves in six railway rails (in nanoseconds, 36 100 subtracted). Three measurements per rail were taken; we have a one-way classification.

The fixed-effects model in cell means form, $y_{ij} = \beta_i + \varepsilon_{ij}$ (where i indicates the rail), can be fitted in R as follows:

```
> rail.lm.1 <- lm(travel ~ Rail - 1, data = rail.df)
> rail.lm.1

Call:
lm(formula = travel ~ Rail - 1, data = rail.df)

Coefficients:
Rail1  Rail2  Rail3  Rail4  Rail5  Rail6
54.00  31.67  84.67  96.00  50.00  82.67
```

We compare this model to the null model, $y_{ij} = \beta + \varepsilon_{ij}$:
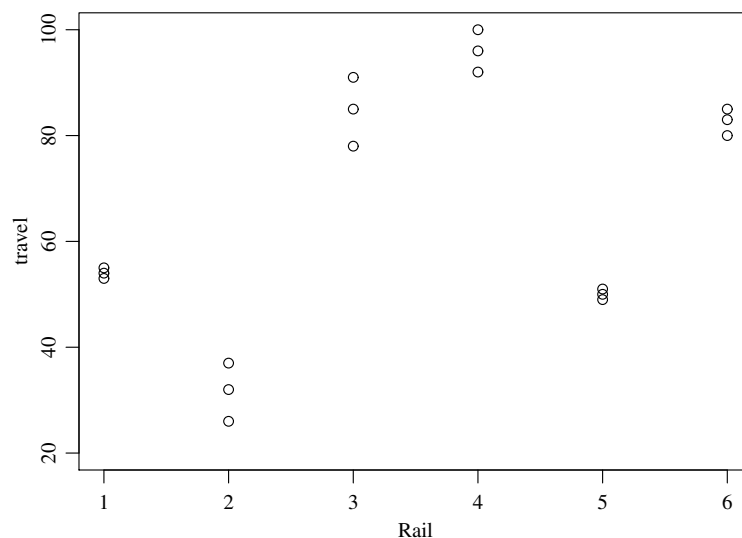
Figure 1.1: The rail data.

```
> rail.lm.0 <- lm(travel ~ 1, data = rail.df)
> anova(rail.lm.0, rail.lm.1)
Analysis of Variance Table

Model 1: travel ~ 1
Model 2: travel ~ Rail - 1
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1     17 9504.5
2     12  194.0  5    9310.5 115.18 1.033e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The rail effect is highly significant. The same F test can be obtained from the output of:

```
> summary(lm(travel ~ Rail, data = rail.df))
[...]

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   54.000      2.321  23.262 2.37e-11 ***
Rail2        -22.333      3.283  -6.803 1.90e-05 ***
Rail3         30.667      3.283   9.341 7.44e-07 ***
Rail4         42.000      3.283  12.793 2.36e-08 ***
Rail5         -4.000      3.283  -1.218    0.246
Rail6         28.667      3.283   8.732 1.52e-06 ***
[...]

Residual standard error: 4.021 on 12 degrees of freedom
Multiple R-squared:  0.9796,    Adjusted R-squared:  0.9711
F-statistic: 115.2 on 5 and 12 DF,  p-value: 1.033e-09
```

$\triangle$

Chapter 2

Linear Mixed-Effects (LME) Models

Example 2.1. Rail data (continued, P&B p. 7ff). Instead of the fixed-effects model $y_{ij} = \beta_i + \varepsilon_{ij}$, we use

$$y_{ij} = \beta + b_i + \varepsilon_{ij}, \tag{P\&B 1.4}$$

where β is the expected travel time across the population of the rails that the sample is taken from. b_i is not a parameter, but a random variable representing the difference between the expected travel time for rail i and β. The b_i's are called *random effects* since they correspond to an experimental unit (the rail) that is randomly selected from a population.

We have to make an assumption not only about the distribution of the error term, but also about that of the random effects:

$$\varepsilon_{ij} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, \sigma^2), \quad b_i \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, \sigma_b^2), \tag{P\&B 1.5}$$

and all these random variables are independent.

Since $\mathrm{Var}(y_{ij}) = \mathrm{Var}(b_i) + \mathrm{Var}(\varepsilon_{ij}) = \sigma_b^2 + \sigma^2$ and $\mathrm{Cov}(y_{i1}, y_{i2}) = \mathrm{Var}(b_i) = \sigma_b^2$, observations from the same rail have correlation $\sigma_b^2/(\sigma_b^2 + \sigma^2)$ under this model.

This model has only three parameters: β, σ_b^2, and σ^2 – irrespective of M, the number of experimental units. (Recall that the fixed-effects model had $M{+}1$ parameters – $\beta_1, \ldots, \beta_M$, σ^2.)

The mixed-effects model for the rail data can be fitted using the `lme` function from the `nlme` R package:

```
rail.lme <- lme(travel ~ 1, random = ~ 1 | Rail, data = rail.df)
```

`travel ~ 1` indicates the response and the fixed effects (in this case, an intercept only), and `random = ~ 1 | Rail` indicates that additionally, a random intercept is fitted with a separate value for each rail.

By `summary(rail.lme)`, we obtain the following output:

```
Linear mixed-effects model fit by REML
 Data: rail.df
      AIC      BIC    logLik
  128.177 130.6766 -61.0885
```

```
Random effects:
 Formula: ~1 | Rail
        (Intercept) Residual
StdDev:    24.80547 4.020779

Fixed effects: travel ~ 1
            Value Std.Error DF  t-value p-value
(Intercept) 66.5   10.17104 12 6.538173       0

Standardized Within-Group Residuals:
        Min          Q1         Med         Q3         Max
-1.61882658 -0.28217671  0.03569328  0.21955784  1.61437744

Number of Observations: 18
Number of Groups: 6
```

$\hat{\beta} = 66.5$ is the overall mean. The standard deviations of the random effect and the error term are estimated as $\hat{\sigma}_b = 24.805$ and $\hat{\sigma} = 4.0208$. (Since the data is balanced, $\hat{\sigma}$ is the same as in the model with a fixed-effect parameter for the expected value of each rail.) Using the `lme4` package, a very similar output can be obtained from the `lmer` function, where fixed and random effects are combined into a single parameter:

`summary(lmer(travel ~ 1 + (1 | Rail), data = rail.df))`   △

Terminology

Fixed effects are model parameters. Typically, they are effects of a variable that we are interested in – we want to estimate them and/or to test hypotheses about the effect, e. g., for specific levels of a factor. If the variable is a factor, it has only a finite number of levels that are usually known at the planning stage, such as the treatment in a clinical study, the sex of a study participant, the type of fertilizer used in an agricultural experiment, etc.

Random effects are random variables with the same realization for observations within a subject or group. They result from selecting, for some subject or group, one of (infinitely) many possible levels of a factor. Such effects are random in the sense that this selection corresponds to randomly drawing from a population, where each element of the population is designated by one level of the factor. We are not particularly interested in the values for specific subjects or groups, and if we do tests, we usually only want to know whether effects of the factor are present or not. Typical examples of random effects are effects that are attributable to a person or a family, an animal or a litter, or a plot of land. In many cases, different levels of the factor would be used in a repetition of the experiment.

A *fixed-effects model* is a model that contains only fixed effects, while a *random-effects model* only contains random effects. A *mixed-effects model* contains both fixed and random effects.

A *single-level model* has only random effects depending on one classification of the observations – in the rail example, this classification divides the measurements into six

classes according to the rail used. A *multi-level model* has more than one variable with random effects, and these variables are *nested*, i.e., they are hierarchically ordered in the sense that each value of one variable (subgroup) only appears within one value of a variable specifying a superior level (group). A two-level model could, e.g., contain the result of a standardized exam for pupils belonging to different classes (level 2), which in turn belong to different schools (level 1).

Note that "level" in "single-/multi-level model" is used for the layers of a hierarchical model, which is not the same as a level of a factor.

The terminology for multi-level models is not used consistently throughout the literature and software: SAS and MLwiN (a specialized multi-level modelling software) number the levels from the random error of each individual observation (level 1) to the level corresponding to the largest groups of observations (level $L + 1$) and call such a model an $(L + 1)$-level model. We (and R) number from the largest groups (level 1) to the smallest (sub-)groups (level L), do not count the individual observations as a level and therefore call the same model an L-level model.

2.1   Single-Level LME Models

The basic LME model with a single level of grouping (M groups) is

$$\begin{aligned}
\boldsymbol{y}_i &= \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \dots, M, \\
\boldsymbol{b}_i &\sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Psi}), \\
\boldsymbol{\varepsilon}_i &\sim \mathrm{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{n_i}).
\end{aligned} \tag{P\&B 2.1}$$

$\boldsymbol{y}_i$ is the n_i-dimensional response vector for the ith group. $\boldsymbol{\beta}$ is the p-dimensional vector of fixed effects, $\boldsymbol{b}_i$ the q-dimensional vector of random effects in group i. $\boldsymbol{X}_i$ is the fixed-effects regressor matrix ($n_i \times p$), $\boldsymbol{Z}_i$ the random-effects regressor matrix ($n_i \times q$), both for group i. $\boldsymbol{\varepsilon}_i$ is the n_i-dimensional within-group error vector, assumed to have i.i.d. normal components with expectation 0 and variance σ^2. (This assumption will be relaxed later.) $\boldsymbol{b}_i$ and $\boldsymbol{\varepsilon}_i$ are assumed to be independent of each other within each group as well as independent across all groups.

The variance-covariance matrix $\boldsymbol{\Psi}$ of the random effects vector $\boldsymbol{b}_i$ has to be symmetric and can be assumed to be positive definite (since the model could be re-expressed in lower dimension using a positive definite matrix if $\boldsymbol{\Psi}$ was only positive semi-definite).

For computational purposes, we sometimes use the *relative precision factor* $\boldsymbol{\Delta}$, which is any matrix (not necessarily unique) satisfying

$$\frac{\boldsymbol{\Psi}^{-1}}{1/\sigma^2} = \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta}.$$

The variance-covariance matrix of $\boldsymbol{y}_i$ depends on σ^2, $\boldsymbol{Z}_i$, and $\boldsymbol{\Psi}$:

$$
\begin{aligned}
\boldsymbol{V}_i = \operatorname{Var}(\boldsymbol{y}_i) &= \operatorname{Var}(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i) \\
&= \boldsymbol{Z}_i\operatorname{Var}(\boldsymbol{b}_i)\boldsymbol{Z}_i^{\mathrm{T}} + \operatorname{Var}(\boldsymbol{\varepsilon}_i) \\
&= \boldsymbol{Z}_i\boldsymbol{\Psi}\boldsymbol{Z}_i^{\mathrm{T}} + \sigma^2\boldsymbol{I}_{n_i}
\end{aligned}
$$

$\boldsymbol{y}_i$ is normally distributed with expectation $\boldsymbol{X}_i\boldsymbol{\beta}$ and variance-covariance matrix $\boldsymbol{V}_i$.

Sometimes, we want to write the model for all groups in a single equation,

$$
\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{b} + \boldsymbol{\varepsilon},
$$

where

$$
\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_M \end{bmatrix}, \quad
\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_M \end{bmatrix}, \quad
\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z}_2 & & \boldsymbol{0} \\ \vdots & & \ddots & \\ \boldsymbol{0} & \boldsymbol{0} & & \boldsymbol{Z}_M \end{bmatrix},
$$

$$
\boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_M \end{bmatrix}, \quad
\boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \vdots \\ \boldsymbol{\varepsilon}_M \end{bmatrix}.
$$

We then have $\boldsymbol{y} \sim \mathrm{N}(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{V})$ with

$$
\boldsymbol{V} = \begin{bmatrix} \boldsymbol{V}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_2 & & \boldsymbol{0} \\ \vdots & & \ddots & \\ \boldsymbol{0} & \boldsymbol{0} & & \boldsymbol{V}_M \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z}_1\boldsymbol{\Psi}\boldsymbol{Z}_1^{\mathrm{T}} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z}_2\boldsymbol{\Psi}\boldsymbol{Z}_2^{\mathrm{T}} & & \boldsymbol{0} \\ \vdots & & \ddots & \\ \boldsymbol{0} & \boldsymbol{0} & & \boldsymbol{Z}_M\boldsymbol{\Psi}\boldsymbol{Z}_M^{\mathrm{T}} \end{bmatrix} + \sigma^2\boldsymbol{I}_N,
$$

where $N = \sum_{i=1}^{M} n_i$ is the total number of observations.

Example 2.2. Rail data (continued, P&B p. 59). The general model (P&B 2.1) can be specialized into model (P&B 1.4) by setting $M = 6$, $n_i = 3$, $i = 1, \ldots, 6$, $p = q = 1$,

$$
\boldsymbol{X}_i = \boldsymbol{Z}_i = \boldsymbol{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad i = 1, \ldots, 6.
$$

Since $\boldsymbol{b}_i = b_i$, the random effect for group i, is a scalar, $\boldsymbol{\Psi} = [\sigma_b^2]$ is also a scalar, as well as the relative precision factor, $\Delta = \sqrt{\sigma^2/\sigma_b^2}$. $\triangle$

Example 2.3. Growth curves (P&B p. 30ff, 59f). Measurements of the distance between two specific points of the skull were taken every two years from an age of 8 up to 14 years on 27 children (16 boys and 11 girls). In a first step, we only use the data for the girls. Their growth curves are given in Figure 2.1.
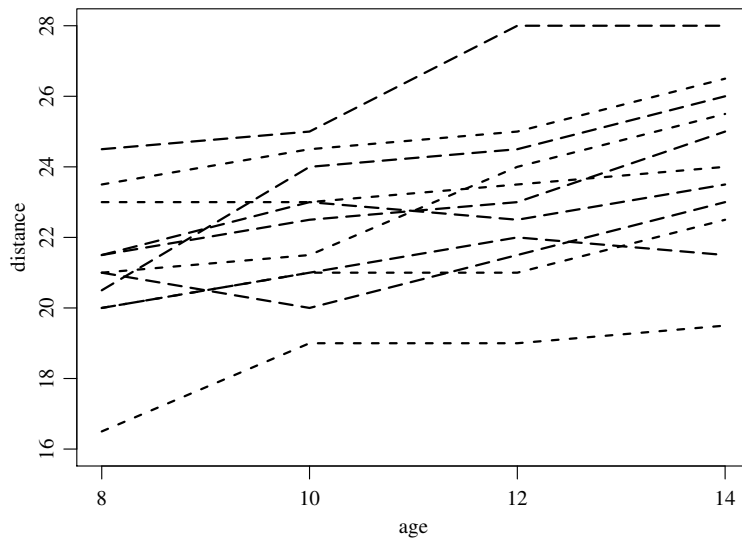
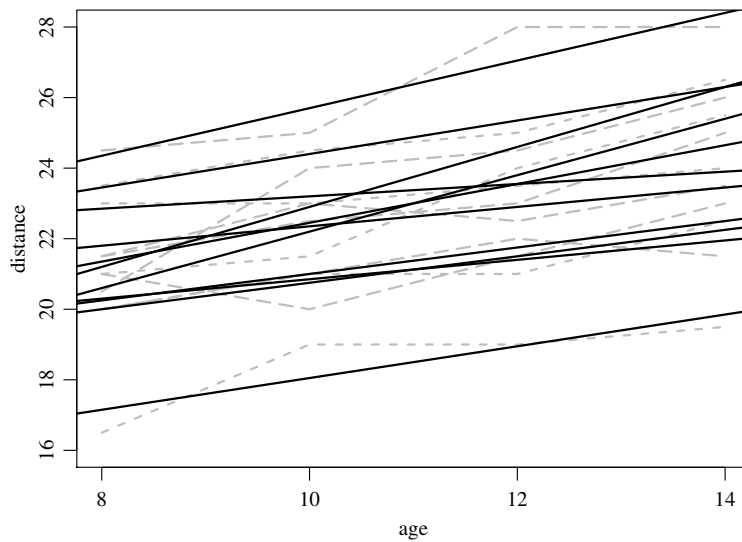Figure 2.1: The growth curves data for 11 girls.

Figure 2.2: Individual regression lines for the 11 girls.

There seems to be a roughly linear increase for most of the girls. Differences in the level as well as in the slope might be relevant.

A simple fixed-effects model approach could be to fit individual regression lines for the 11 girls, estimating a total of 33 parameters (11 intercepts, 11 slopes, and 11 error variances). In R, this can be done using the following call:

```
ortho.fem.lmlist.1 <- lmList(distance ~ age | Subject, data=ortho.fem)
```

The fitted lines are shown in Figure 2.2.

Note that individual regression lines could also be obtained by a single linear fixed-effects model with main effects for age and subject (as a factor) and with an age × subject interaction (assuming the variance to be the same for all girls). This model has 22 parameters for the fixed effects (corresponding to the 11 intercepts and 11 slopes) and a variance parameter.

Since we are not really interested in individual regression lines for these specific girls, but rather in general results, a mixed-effects model approach seems more sensible. This will also help to reduce the number of parameters to be estimated.

In the rails example, we have only looked at a *random intercept model*, i.e., a linear mixed-effects model with a random intercept per observational unit, but no other random effects. Such a model with random-effects regressor matrices

$$\boldsymbol{Z}_i = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

can be fitted here by:

```
ortho.fem.lme.1 <- lme(distance ~ age, random = ~ 1 | Subject, data=ortho.fem)
```

However, in order to have something similar to the individual regression lines from above, we need a linear mixed-effects model with random intercepts and slopes. The corresponding random-effects regressor matrices in our example are

$$\boldsymbol{Z}_i = \begin{bmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \end{bmatrix}.$$

Note that $\boldsymbol{Z}_i = \boldsymbol{X}_i$ and that we still need the fixed effects for the intercept and the slope since the random effects are assumed to be centered at 0.

This model only has 6 parameters – two for the fixed effects (intercept and slope), an error variance, and three parameters for $\boldsymbol{\Psi}$. It can be fitted by:

```
ortho.fem.lme.2 <- lme(distance ~ age, random = ~ age | Subject, data=ortho.fem)
```

We will come back to the comparison of these two models later.

Using the `lmer` function, the corresponding formula objects would be `distance ~ age + (1 | Subject)` and `distance ~ age + (age | Subject)`, respectively. △

Example 2.4. Machines data (P&B p. 21–23, 27–29). In this example, the productivity of six workers was measured three times on each of three different machines – see Figure 2.3. There seem to be differences between machines and probably also between workers.

The workers are a random sample of possible persons. The three machines, on the other hand, are the only three types that are available. Therefore, fixed effects will be used for the machines and random effects for the workers.
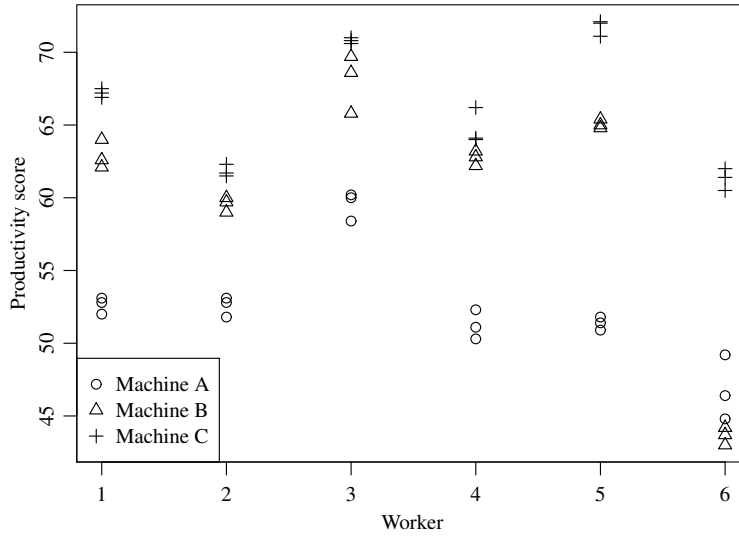
Figure 2.3: Productivity scores for 6 workers on 3 machines.

If y_{ijk} denotes the productivity score for the kth replication of the ith worker on machine j, let

$$
\boldsymbol{y}_i = \begin{bmatrix} y_{i11} \\ y_{i12} \\ y_{i13} \\ y_{i21} \\ y_{i22} \\ y_{i23} \\ y_{i31} \\ y_{i32} \\ y_{i33} \end{bmatrix}.
$$

A model with fixed effects for `Machine` and random effects for `Worker` (without interaction) can be specified based on the general model (P&B 2.1) using

$$
\boldsymbol{X}_i = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{Z}_i = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},
$$

a parameter vector $\boldsymbol{\beta} \in \mathbb{R}^3$ for the fixed intercept and machine effects (using the first machine as the reference), and random intercepts $\boldsymbol{b}_i = b_i \in \mathbb{R}^1$ for the different workers.
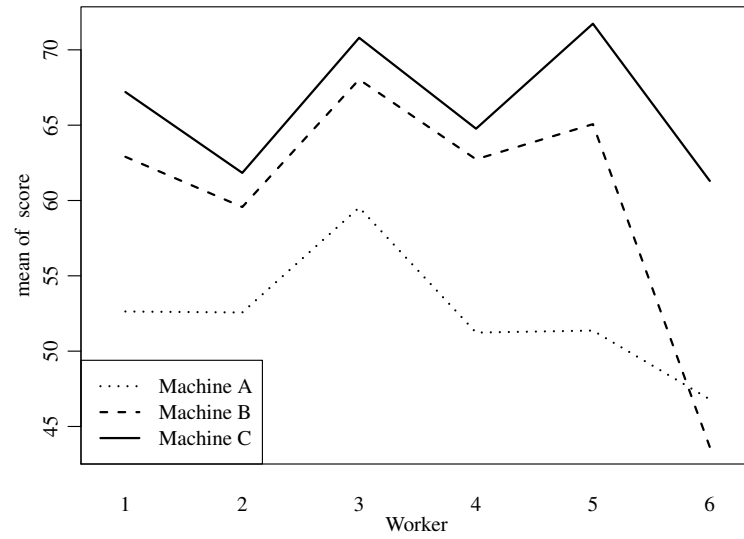
Figure 2.4: Interaction plot for the machines data.

This model can be fitted using:

```
machines.lme.1 <- lme(score ~ Machine, random = ~ 1 | Worker, data=machines.df)
```

(Since a different parametrization for the fixed effects is used, the results in P&B are different from those obtained in R using this code. The results from the book can be reproduced using the additional parameter `contrasts=list(Machine=contr.helmert)`.)

```
with(machines.df, interaction.plot(Worker, Machine, score))
```

produces an interaction plot (Figure 2.4), which shows that some interaction term for `Worker` $\times$ `Machine` should probably be included. An interaction of a random-effects and a fixed-effects term is again random. One possibility for including such a term is to use separate random-effects terms for the three machines. To do so, the previous model can be modified as follows:

$$\boldsymbol{Z}_i = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$\boldsymbol{b}_i \in \mathbb{R}^3$ is now the vector of the effects of the three machines for worker i.

This more flexible model can be fitted using:

```
> machines.lme.3 <- lme(score ~ Machine, random = ~ Machine - 1 | Worker,
+   data=machines.df)
> summary(machines.lme.3)
Linear mixed-effects model fit by REML
 Data: machines.df
       AIC      BIC    logLik
  228.3112 247.6295 -104.1556

Random effects:
 Formula: ~Machine - 1 | Worker
 Structure: General positive-definite, Log-Cholesky parametrization
         StdDev    Corr
MachineA 4.0792807 MachnA MachnB
MachineB 8.6252908 0.803
MachineC 4.3894795 0.623  0.771
Residual 0.9615766

Fixed effects: score ~ Machine
               Value Std.Error DF   t-value p-value
(Intercept) 52.35556  1.680711 46 31.150834  0.0000
MachineB     7.96667  2.420851 46  3.290854  0.0019
MachineC    13.91667  1.540100 46  9.036211  0.0000
 Correlation:
        (Intr) MachnB
MachineB  0.463
MachineC -0.374  0.301

Standardized Within-Group Residuals:
        Min          Q1         Med         Q3         Max
-2.39354008 -0.51377575  0.02690829  0.47245472  2.53338699

Number of Observations: 54
Number of Groups: 6
```

Ψ, the variance-covariance matrix of the b_i's, may be any positive-definite, symmetric 3×3 matrix. The variances for the random effects of the different machines may therefore be different. As can be seen from the output of summary(machines.lme.3), the estimated standard deviations differ by a factor of more than 2. The estimated correlations between the random effects of the three machines are between 0.623 and 0.803 – since we expect some effect of the worker, it is plausible that they are all positive.

We will see a model with a less general structure for a Worker × Machine interaction later.   △