Peter von Rohr
Institute of Agricultural Sciences
D-USYS
ETH Zurich

751-7602-00 V
Solutions for Exam in
Applied Statistical Methods
in Animal Sciences
Spring Semester 2023

Date: 2023-05-22

Name:

Legi-Nr:

Problem	Maximum Number of Points	Number of Points Reached
1	23	
2	12	
3	50	
4	40	
5	16	
Total	141	

Questions in German are in italics

Problem 1: Linear Regression

Given is the following dataset on methane yield (in g/kg of Dry Matter Intake) and the nonvolatile milk metabolite acetone (as NMR-based relative area) for 12 cows. The dataset is available under the link shown below the table

Gegeben ist der nachfolgend gezeigte Datensatz mit Methan (in g/kg Trockensubstanzaufnahme) und dem nicht-flüchtigen Milkmetaboliten Aceton (als relative Fläche aus der NMR) für 12 Kühe. Der Datensatz ist verfügbar unter dem Link, welcher unter der Tabelle gezeigt wird.

Cow	Acetone	Methane
1	0.0276	24.41
2	0.0316	24.03
3	0.0807	22.01
4	0.0325	24.83
5	0.0172	24.70
6	0.0388	23.77
7	0.0276	23.98
8	0.0437	23.37
9	0.0322	24.33
10	0.0353	24.05
11	0.0631	22.23
12	0.0081	25.34

```
## https://charlotte-ngs.github.io/asmss2023/data/asm_exam_p01.csv
```

a) Fit a linear regression model with methane yield as response variable and acetone as predictor variable. What are the estimated values for

- the slope,
- the intercept and
- the standard deviation of the residuals

based on fitted regression model?

Passen Sie ein lineares Regressionsmodell mit Methan als Zielgrösse und Aceton als unabhängige Variable (Predictor). Wie gross sind die aufgrund des Regressionsmodells geschätzten Werte für

- *die Steigung,*
- *der Achsenabschnitt und*
- *die Standardabweichung der Residuen*

4

Solution

The data is read as given by

```
s_data_url_root <- "https://charlotte-ngs.github.io/asmss2023/data"
s_data_url_p1 <- file.path(s_data_url_root, "asm_exam_p01.csv")
tbl_p1 <- readr::read_delim(file = s_data_url_p1, delim = ",")
```

The linear regression model is fit by

```
lm_p1 <- lm(formula = Methane ~ Acetone, data = tbl_p1)
smry_p1 <- summary(lm_p1)
smry_p1
```

```
##
## Call:
## lm(formula = Methane ~ Acetone, data = tbl_p1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.39949 -0.17106 -0.00179  0.10155  0.71312
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  25.6966     0.2018 127.366  < 2e-16 ***
## Acetone     -48.6078     4.9257  -9.868  1.8e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.316 on 10 degrees of freedom
## Multiple R-squared:  0.9069, Adjusted R-squared:  0.8976
## F-statistic: 97.38 on 1 and 10 DF,  p-value: 1.795e-06
```

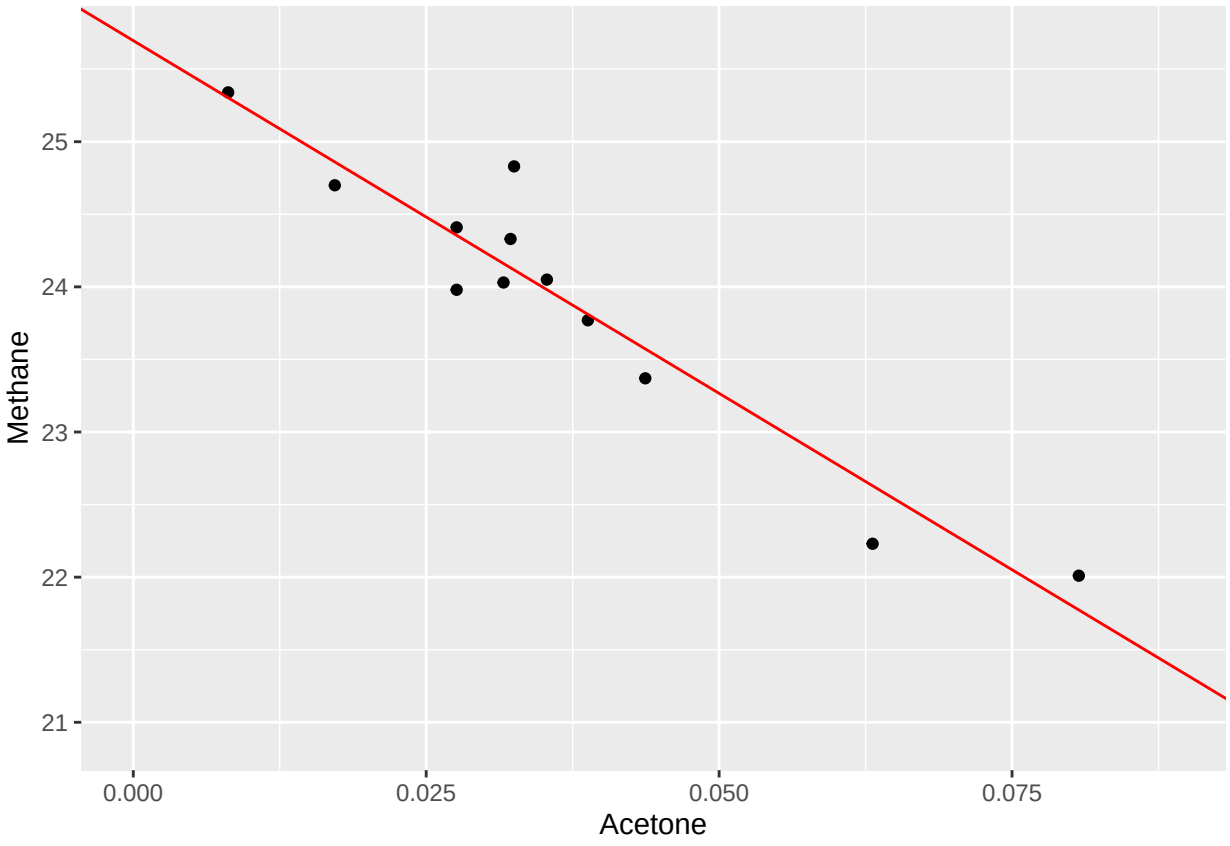
The estimates for the quantities asked in the problem are listed in the following table

Quantity	Estimate
Slope	-48.6077793
Intercept	25.6966375
SD of Residuals	0.3160059

b) The plot below shows a scatter-plot of the given data on Acetone and Methane together with the fitted regression line. Please enter into the plot the slope, the intercept and all residuals.

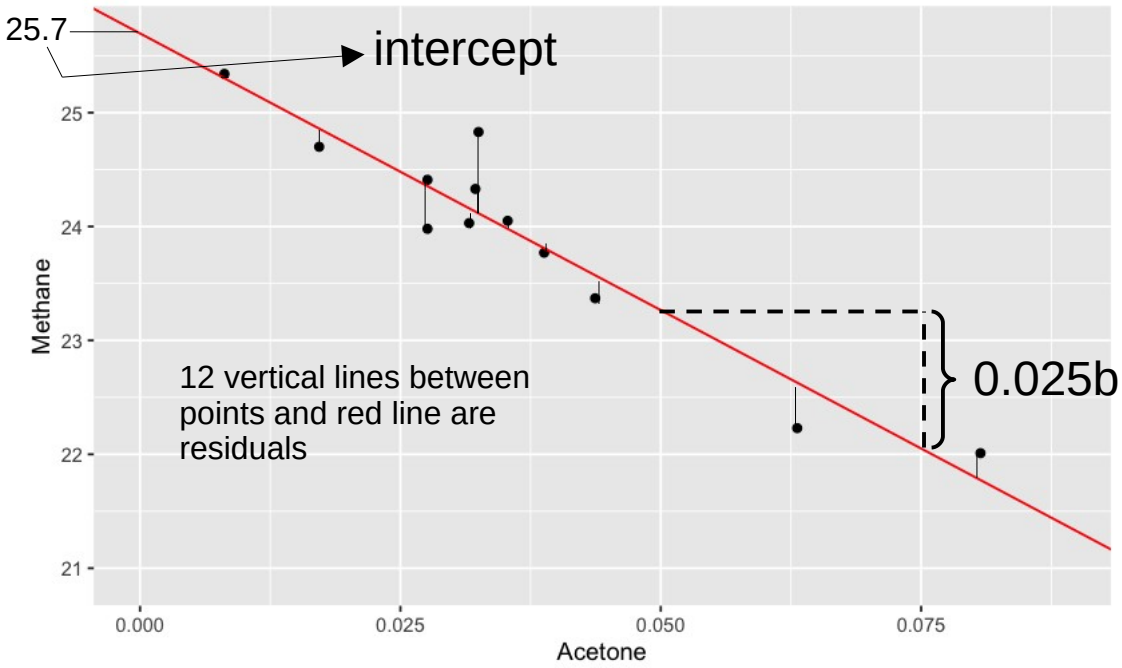
Die unten gezeigte Grafik zeigt den Scatterplot der gegebenen Daten zu Aceton und Methan zusammen mit der angepassten Regressionslinie. Bitte zeichnen Sie in den gegebenen Plot die folgenden Grössen ein: Steigung, Achsenabschnitt und alle Residuen

14



Solution

The required quantities entered in the plot are shown below


25.7
intercept
25
24
23
22
21
Methane
12 vertical lines between points and red line are residuals
0.025b
0.000
0.025
0.050
0.075
Acetone

c) The table below contains acetone measurements for 3 additional cows. Predict their methane values based on the linear regression that you found in Problem 1a). If you did not manage to solve Problem 1a) you can assume an intercept of 25 and a slope of -40 for the prediction of methane based on acetone. Please specify whether all of the predicted methane values are valid or whether there is a problem with some of the values. If you find a problem please specify which values have a problem and what the problem is.

Die nachfolgende Tabelle enthält Acetonwerte für 3 zusätzliche Kühe. Berechnen Sie die geschätzten Methanwerte für diese Kühe aufgrund der geschätzten Regression aus Aufgabe 1a). Falls Sie die Aufgabe 1a) nicht gelöst haben, dann können Sie einen Achsenabschnitt von 25 und eine Steigung von -40 für die Vorhersage der Methanwerte annehmen. Bitte geben Sie an, ob die geschätzten Methanwerte alle gültig sind oder ob es mit gewissen Werten ein Problem gibt. Falls Sie Probleme finden, geben Sie bitte an welche Werte problematisch sind und was genau das Problem ist.

5

Cows with additional acetone values

Cow	Acetone
13	0.0532
14	0.0120
15	0.1090

Solution

The predicted values with results from Problem 1a)

```
tbl_pred_meth <- tbl_added_ace
tbl_pred_meth$Methane <- lm_p1$coefficients[["(Intercept)"]] +
  lm_p1$coefficients[["Acetone"]] * tbl_pred_meth$Acetone
knitr::kable(tbl_pred_meth,
             booktabs = TRUE,
             longtable = TRUE)
```

Cow	Acetone	Methane
13	0.0532	23.11070
14	0.0120	25.11334
15	0.1090	20.39839

The predicted values with assumed values

```
tbl_pred_meth <- tbl_added_ace
tbl_pred_meth$Methane <- n_intercept_value +
  n_slope_value * tbl_pred_meth$Acetone
knitr::kable(tbl_pred_meth,
             booktabs = TRUE,
             longtable = TRUE)
```

Cow	Acetone	Methane
13	0.0532	22.872
14	0.0120	24.520
15	0.1090	20.640

Find the cow with acetone outside of valid range

```
vec_upper <- which(tbl_added_ace$Acetone > max(tbl_p1$Acetone))
vec_lower <- which(tbl_added_ace$Acetone < min(tbl_p1$Acetone))
vec_outside <- union(vec_upper,vec_lower)
vec_outside_cow_id <- vec_outside + n_nr_cow
```

The problem is that the acetone value for cow 15 are outside of the valid range of acetone measurements and hence the predictions are not valid.

Problem 2: Fixed Linear Effects Model

The dataset from Problem 1 is extended by the breed of every cow. The extended dataset is shown in the table below. The extended dataset can be read from the link shown below the table.

Der Datensatz von Aufgabe 1 wurde um die Rasse der Kühe erweitert. Der erweiterte Datensatz ist in der nachfolgenden Tabelle gezeigt. Der erweiterte Datensatz kann vom nachfolgenden Link gelesen werden.

Cow	Acetone	Methane	Breed
1	0.0276	24.41	Angus
2	0.0316	24.03	Angus
3	0.0807	22.01	Limousin
4	0.0325	24.83	Limousin
5	0.0172	24.70	Angus
6	0.0388	23.77	Limousin
7	0.0276	23.98	Angus
8	0.0437	23.37	Limousin
9	0.0322	24.33	Limousin
10	0.0353	24.05	Limousin
11	0.0631	22.23	Limousin
12	0.0081	25.34	Angus

```
## https://charlotte-ngs.github.io/asmss2023/data/asm_exam_p02.csv
```

a) Fit a fixed linear effects model with methane as the response and with acetone and breed as predictor variables.

Passen Sie ein fixes lineares Modell an mit Methan als Zielgrösse und mit Aceton und Rasse als unabhängige Variable.

2

Solution

The fixed linear effect model

```
lm_p2 <- lm(Methane ~ Acetone + Breed, data = tbl_p2)
smry_lm_p2 <- summary(lm_p2)
smry_lm_p2
```

```
##
## Call:
## lm(formula = Methane ~ Acetone + Breed, data = tbl_p2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38642 -0.18341 -0.02125  0.10191  0.54966
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    25.7111     0.1912 134.442 3.54e-16 ***
## Acetone       -54.3766     6.0956  -8.921 9.18e-06 ***
```

```
## BreedLimousin   0.3365     0.2290   1.469    0.176
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2991 on 9 degrees of freedom
## Multiple R-squared:  0.9249, Adjusted R-squared:  0.9082
## F-statistic: 55.41 on 2 and 9 DF,  p-value: 8.722e-06
```

b) Fit a fixed linear effects model with methane as response variable and the predictors acetone and breed, including an interaction term between acetone and breed.

Passen Sie ein fixes lineares Modell an mit Methan als Zielgrösse und mit den unabhängigen Variablen Aceton und Rasse, inklusive eines Interaktionsterms zwischen Aceton und Rasse.

2

Solution

The fixed linear effects model with interaction

```
lm_p2_inta <- lm(Methane ~ Acetone * Breed, data = tbl_p2)
smry_lm_p2_inta <- summary(lm_p2_inta)
smry_lm_p2_inta
```

```
##
## Call:
## lm(formula = Methane ~ Acetone * Breed, data = tbl_p2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38977 -0.18077 -0.01976  0.09134  0.55253
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            25.7362     0.3959  65.014 3.49e-12 ***
## Acetone               -55.4958    16.4840  -3.367  0.00983 **
## BreedLimousin           0.3019     0.5276   0.572  0.58289
## Acetone:BreedLimousin   1.3225    17.9188   0.074  0.94298
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3172 on 8 degrees of freedom
## Multiple R-squared:  0.9249, Adjusted R-squared:  0.8968
## F-statistic: 32.86 on 3 and 8 DF,  p-value: 7.572e-05
```

c) What are the expected methane values for the cows of a given breed and with given acetone values, using the model with and without interactions.

Welches sind die erwarteten Methanwerte für die unten aufgelisteten Kühe einer bestimmten Rassen und mit beobachteten Acetonwerten. Unterscheiden Sie dabei nach den Modellen mit und ohne Interaktion

Cow	Acetone	Breed	Expected Methane without Interaction	Expected Methane with Interaction
13	0.0532	Limousin		
14	0.0120	Limousin		
15	0.0090	Angus		
16	0.0169	Angus		

Solution

The computation is done separately per breed

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
tbl_added_ace_li <- tbl_added_ace_breed |>
  filter(Breed == "Limousin")
tbl_added_ace_li$`Expected Methane without Interaction` <-
  smry_lm_p2$coefficients["(Intercept)", "Estimate"] +
  smry_lm_p2$coefficients["Acetone", "Estimate"] * tbl_added_ace_li$Acetone +
  smry_lm_p2$coefficients["BreedLimousin", "Estimate"]
tbl_added_ace_li$`Expected Methane with Interaction` <-
 smry_lm_p2_inta$coefficients["(Intercept)", "Estimate"] +
 smry_lm_p2_inta$coefficients["Acetone", "Estimate"] * tbl_added_ace_li$Acetone +
 smry_lm_p2_inta$coefficients["BreedLimousin", "Estimate"] +
 smry_lm_p2_inta$coefficients["Acetone:BreedLimousin", "Estimate"] * tbl_added_ace_li$Acetone
knitr::kable(tbl_added_ace_li,
             booktabs = TRUE,
             longtable = TRUE)
```

Cow	Acetone	Breed	Expected Methane without Interaction	Expected Methane with Interaction
13	0.0532	Limousin	23.15475	23.15609

Cow	Acetone	Breed	Expected Methane without Interaction	Expected Methane with Interaction
14	0.0120	Limousin	25.39506	25.38803

The same for Angus animals

```
tbl_added_ace_an <- tbl_added_ace_breed |>
  filter(Breed == "Angus")
tbl_added_ace_an$`Expected Methane without Interaction` <-
  smry_lm_p2$coefficients["(Intercept)", "Estimate"] +
  smry_lm_p2$coefficients["Acetone", "Estimate"] * tbl_added_ace_an$Acetone
tbl_added_ace_an$`Expected Methane with Interaction` <-
  smry_lm_p2_inta$coefficients["(Intercept)", "Estimate"] +
  smry_lm_p2_inta$coefficients["Acetone", "Estimate"] * tbl_added_ace_an$Acetone
knitr::kable(tbl_added_ace_an,
             booktabs = TRUE,
             longtable = TRUE)
```

Cow	Acetone	Breed	Expected Methane without Interaction	Expected Methane with Interaction
15	0.0090	Angus	25.22173	25.23675
16	0.0169	Angus	24.79216	24.79834

Problem 3: Pedigree BLUP

Use the data shown in the table below to fit linear mixed effects models. The column `P` corresponds to values of phenotypic observations and is to be used as response variable in the linear model. The column `Sex` contains the sex of the animal and is to be used as a fixed effect in the linear model.

Verwenden Sie die Daten in der unten gezeigten Tabelle um ein lineares gemischtes Modell anzupassen. Die Kolonne mit dem Titel 'P' enthält Werte von phänotypischen Beobachtungen, welche als Zielgrössen im Modell verwendet werden sollen. Die Kolonne 'Sex' enthält das Geschlecht der Tiere und soll als ein fixer Effekt ins gemischte Modell einfliessen.

ID	SIRE	DAM	SEX	P
7	1	4	f	65.7
8	2	5	f	61.6
9	3	6	m	44.3
10	9	8	m	43.2
11	1	7	f	67.9
12	10	11	f	68.2

```
## https://charlotte-ngs.github.io/asmss2023/data/asm_exam_p03.csv
```

a) Use a sire model to predict breeding values for all sires. Please specify the model in the form of an equation and explain the meaning of all model components. Write down the expected values and the variance-covariance matrices for all random effects in the model. Put all the information from the data into the model components consisting of response variable, fixed effects, random effects, and their associated design matrices. Construct the mixed model equations and provide solutions to these equations. Extract predicted breeding values from these solutions.

   The following variance components can be assumed to be given as

   - sire variance $\sigma_s^2 = 6.25$
   - residual variance $\sigma_e^2 = 118.75$

   Verwenden Sie ein Vatermodell zur Schätzung der Zuchtwerte aller männlichen Tiere. Bitte geben Sie die Modellgleichung and und erläutern Sie die Bedeutung aller Modellkomponenten. Schreiben Sie die Erwartungswerte und die Varianz-Kovarianzmatrizen aller zufälligen Effekte im Modell auf. Verwenden Sie alle Informationen aus den Daten für die genaue Spezifikation der Modellkomponenten, wie zum Beispiel Zielgrösse, fixe Effekte, zufällige Effekte und die zugehörigen Designmatrizen. Stellen Sie die Mixed-Model-Gleichungen auf und lösen Sie diese Gleichungen. Extrahieren Sie die Zuchtwerte aller männlichen Tiere aus den Lösungen.

   Die folgenden Varianzkomponenten können als gegeben angenommen werden:

   - Sire Varianz $\sigma_s^2 = 6.25$
   - Varianz der Residuen $\sigma_e^2 = 118.75$

25

Solution

- Model as equation:

$$y = Xb + Zs + e$$

where

y	vector of observations
b	vector of fixed effects
s	vector of random sire breeding values
e	vector of random residuals
X	design matrix linking observations to fixed effects
Z	design matrix linking observations to sire breeding values

- Expected values and the variance-covariance matrices:

Random effects are y, s and e. Their expected values correspond to

$$E \begin{bmatrix} y \\ s \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \end{bmatrix}$$

The variance-covariance matrices

$$var \begin{bmatrix} y \\ s \\ e \end{bmatrix} = \begin{bmatrix} V & ZS & R \\ SZ^T & S & 0 \\ R & 0 & R \end{bmatrix}$$

where $var(s) = S = A_s * \sigma_s^2$, $var(e) = R = I * \sigma_e^2$ and $var(y) = V = ZSZ^T + R$ with given sire-variance component σ_s^2 and given residual variance component σ_e^2.

- Information from data to model components

$$y = \begin{bmatrix} 65.7 \\ 61.6 \\ 44.3 \\ 43.2 \\ 67.9 \\ 68.2 \end{bmatrix}, b = \begin{bmatrix} b_f \\ b_m \end{bmatrix}, s = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_9 \\ s_{10} \end{bmatrix}, e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix}$$

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

- Mixed model equations:

$$\begin{bmatrix} X^TX & X^TZ \\ Z^TX & Z^TZ + \lambda_s * A_s^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} X^Ty \\ Z^Ty \end{bmatrix}$$

```
## Loading required package: lme4
```

```
## Loading required package: Matrix
```

```
## as(<dtTMatrix>, "dtCMatrix") is deprecated since Matrix 1.5-0; do as(., "CsparseMatrix") instead
```

where $\lambda_s = \frac{\sigma_e^2}{\sigma_s^2} = \frac{118.75}{6.25} = 19$ and A_s corresponds to the sire relationship matrix given by

$$A_s = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0.5 & 0.25 \\ 0 & 0 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0.25 & 0.5 & 1 \end{bmatrix}$$

The inverse A_s^{-1} is

$$A_s^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1.33333333333333 & -0.666666666666667 & 0 \\ 0 & 0 & -0.666666666666667 & 1.66666666666667 & -0.666666666666667 \\ 0 & 0 & 0 & -0.666666666666667 & 1.33333333333333 \end{bmatrix}$$

- Solutions of MME

```
mat_xtx <- crossprod(mat_X_sex)
mat_xtz <- crossprod(mat_X_sex, mat_Z_sire)
mat_ztx <- t(mat_xtz)
mat_ztz_lAsinv <- crossprod(mat_Z_sire) + lambda_s * mat_As_inv
mat_coef_sire <- rbind(cbind(mat_xtx, mat_xtz), cbind(mat_ztx, mat_ztz_lAsinv))
mat_rhs_sire <- rbind(crossprod(mat_X_sex, vec_y), crossprod(mat_Z_sire, vec_y))
mat_sol_sire <- solve(mat_coef_sire, mat_rhs_sire)
mat_sol_sire
```

```
##          [,1]
##   65.82868389
##   43.70540204
## 1  0.09250630
## 2 -0.21143419
## 3  0.04421572
## 4  0.04498021
## 5  0.11168603
```

- Predicted sire breeding values: The first two elements are solutions of fixed effects, the remaining elements are predicted breeding values

Sire	Predicted Breeding Values
1	0.0925063
2	-0.2114342
3	0.0442157
9	0.0449802
10	0.1116860

b) Use an animal model to predict breeding values for all animals in the above shown dataset. Please specify the model in the form of an equation and explain the meaning of all model components. Write down the expected values and the variance-covariance matrices for all random effects in the model. Put all the information from the data into the model components consisting of response variable, fixed effects, random effects, and their associated design matrices. Construct the mixed model equations and provide solutions to these equations. Extract predicted breeding values from these solutions.

The following variance components can be assumed to be given as

- additive genetic variance $\sigma_u^2 = 25$
- residual variance $\sigma_e^2 = 100$

Verwenden Sie ein Tiermodell für die Schätzung der Zuchtwerte aller Tiere im oben gegebenen Datensatz.Bitte geben Sie die Modellgleichung and und erläutern Sie die Bedeutung aller Modellkomponenten. Schreiben Sie die Erwartungswerte und die Varianz-Kovarianzmatrizen aller zufälligen Effekte im Modell auf. Verwenden Sie alle Informationen aus den Daten für die genaue Spezifikation der Modellkomponenten, wie zum Beispiel Zielgrösse, fixe Effekte, zufällige Effekte und die zugehörigen Designmatrizen. Stellen Sie die Mixed-Model-Gleichungen auf und lösen Sie diese Gleichungen. Extrahieren Sie die Zuchtwerte aller männlichen Tiere aus den Lösungen.

Die folgenden Varianzkomponenten können als gegeben angenommen werden:

- Additive Genetische Varianz $\sigma_u^2 = 25$
- Varianz der Residuen $\sigma_e^2 = 100$

25

Solution

- Model as equation:

$$y = Xb + Zu + e$$

where

y	vector of observations
b	vector of fixed effects
u	vector of random breeding values
e	vector of random residuals
X	design matrix linking observations to fixed effects
Z	design matrix linking observations to sire breeding values

- Expected values and the variance-covariance matrices:

Random effects are y, u and e. Their expected values correspond to

$$E \begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \end{bmatrix}$$

The variance-covariance matrices

$$var \begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} V & ZU & R \\ UZ^T & U & 0 \\ R & 0 & R \end{bmatrix}$$

where $var(u) = U = A * \sigma_u^2$, $var(e) = R = I * \sigma_e^2$ and $var(y) = V = ZUZ^T + R$ with given additive genetic variance component σ_u^2 and given residual variance component σ_e^2.

- Information from data to model components

$$y = \begin{bmatrix} 65.7 \\ 61.6 \\ 44.3 \\ 43.2 \\ 67.9 \\ 68.2 \end{bmatrix}, b = \begin{bmatrix} b_f \\ b_m \end{bmatrix}, s = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \\ u_9 \\ u_{10} \\ u_{11} \\ u_{12} \end{bmatrix}, e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix}$$

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

$$Z =$$

- Mixed model equations:

$$\begin{bmatrix} X^TX & X^TZ \\ Z^TX & Z^TZ + \lambda_u * A^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X^Ty \\ Z^Ty \end{bmatrix}$$

where $\lambda_u = \frac{\sigma_e^2}{\sigma_u^2} = \frac{100}{25} = 4$ and A corresponds to the numerator relationship matrix given by

$$A =$$

The inverse A_s^{-1} is

$$A^{-1} =$$

- Solutions of MME

```
mat_xtx <- crossprod(mat_X_sex)
mat_xtz <- crossprod(mat_X_sex, mat_Z_ani)
mat_ztx <- t(mat_xtz)
mat_ztz_lAinv <- crossprod(mat_Z_ani) + lambda_u * mat_A_inv
mat_coef_ani <- rbind(cbind(mat_xtx, mat_xtz), cbind(mat_ztx, mat_ztz_lAinv))
mat_rhs_ani <- rbind(crossprod(mat_X_sex, vec_y), crossprod(mat_Z_ani, vec_y))
mat_sol_ani <- solve(mat_coef_ani, mat_rhs_ani)
mat_sol_ani
```

```
##           [,1]
##    65.62320543
##    43.76069918
## 1   0.44172160
## 2  -0.36860794
## 3   0.08764269
## 4   0.12020890
## 5  -0.36860794
## 6   0.08764269
## 7   0.40117415
## 8  -0.73721587
## 9   0.17528538
## 10 -0.19668373
## 11  0.74296057
## 12  0.50025945
```

- Predicted breeding values for all animals: The first two elements are solutions of fixed effects, the remaining elements are predicted breeding values

Animals	Predicted Breeding Values
1	0.4417216
2	-0.3686079
3	0.0876427
4	0.1202089
5	-0.3686079
6	0.0876427
7	0.4011741
8	-0.7372159
9	0.1752854
10	-0.1966837
11	0.7429606
12	0.5002594

Problem 4: Genomic BLUP

Use the dataset shown in the table below to predict genomic breeding values. The column `P` contains phenotypic observations which are to be used as response variables. The column `SEX` contains the sex of the animal which is to be used as fixed effect. The columns `SNP1` to `SNP5` contain marker genotypes where the number stands for the number of favourite alleles at the respective marker position.

Verwenden Sie den nachfolgend gezeigten Datensatz für die Schätzung von genomischen Zuchtwerten. Die Kolonne 'P' enhält phänotypische Beobachtungen, welche als Zielgrössen zu verwenden sind. Die Kolonne 'SEX' enthält das Geschlecht der Tiere, welche als fixe Effekte verwendet werden. Die Kolonnen 'SNP1' bis 'SNP5' enthalten Markergenotypen, wobei der Wert der Anzahl an positiven Allelen an der entsprechenden Markerposition entsprechen.

ID	SEX	P	SNP1	SNP2	SNP3	SNP4	SNP5
7	m	115.4	2	1	1	1	0
8	f	84.4	2	0	2	1	0
9	f	86.8	0	0	2	1	1
10	m	76.3	1	0	2	1	0
11	f	122.3	2	1	1	2	0
12	f	115.8	2	1	0	2	0

```
## https://charlotte-ngs.github.io/asmss2023/data/asm_exam_p04.csv
```

a) Use a marker effect model in a two step procedure to predict genomic breeding values for all animals shown in the above dataset. The ratio between residual variance and variance of the marker effects is 2.2.

Verwenden Sie ein Markereffektmodell in einem Zwei-Schritt-Verfahren um genomische Zuchtwerte für alle Tiere im oben gezeigten Datensatz zu schätzen.Das Verhältnis zwischen Restvarianz und Markervarianz beträgt 2.2.

20

Solution

The two step procedure consists of

1. marker effect estimates
2. prediction of breeding values

- Marker effects can in principle be estimated using `lm()` in a regression model. But the dataset is too small to yield valid estimates. Hence, we have to use a mixed model.

$$y = Xb + Wq + e$$

with y the vector of observations, b the vector of fixed effects, q the vector of random marker effects, e the vector of random residuals and design matrices X and W. Inserting information from the data leads to

```r
# vector y
vec_y <- tbl_p4$P
n_nr_rec_p4 <- length(vec_y)
# matrix X
mat_X <- model.matrix(lm(P ~ 0 + SEX, data = tbl_p4))
attr(mat_X, "assign") <- NULL
attr(mat_X, "contrasts") <- NULL
colnames(mat_X) <- NULL
# vector b
vec_b <- c("b_{f}", "b_{m}")
# martix W
tbl_geno <- tbl_p4 %>% select(SNP1, SNP2, SNP3, SNP4, SNP5)
mat_W <- as.matrix(tbl_geno)
# vector q
vec_q <- sapply(1:ncol(mat_W), function(x) paste("q_", x, sep = ""))
# vector e
vec_e <- sapply(1:n_nr_rec_p4, function(x) paste("e_", x, sep = ""))
```

$$y = \begin{bmatrix} 115.4 \\ 84.4 \\ 86.8 \\ 76.3 \\ 122.3 \\ 115.8 \end{bmatrix}, X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}, b = \begin{bmatrix} b_f \\ b_m \end{bmatrix}, W = \begin{bmatrix} 2 & 1 & 1 & 1 & 0 \\ 2 & 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 & 1 \\ 1 & 0 & 2 & 1 & 0 \\ 2 & 1 & 1 & 2 & 0 \\ 2 & 1 & 0 & 2 & 0 \end{bmatrix}, q = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \end{bmatrix}, e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix}$$

$$E \begin{bmatrix} y \\ q \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \end{bmatrix}$$

The variance-covariance matrices

$$var \begin{bmatrix} y \\ q \\ e \end{bmatrix} = \begin{bmatrix} ZQZ^T + R & ZQ & R \\ QZ^T & Q & 0 \\ R & 0 & R \end{bmatrix}$$

with

- $var(e) = R = I * \sigma_e^2$,
- $var(q) = Q = I * \sigma_q^2$
- σ_e^2 the residual variance component
- σ_q^2 the marker variance component

Setting up mixed model equations

```r
# coefficient matrix
mat_xtx <- crossprod(mat_X)
mat_xtw <- crossprod(mat_X, mat_W)
mat_wtx <- t(mat_xtw)
mat_wtw <- crossprod(mat_W)
mat_wtw_qinv_lambda <- mat_wtw + lambda_q * diag(1,nrow = nrow(mat_wtw))
mat_coef_marker <- rbind(cbind(mat_xtx, mat_xtw), cbind(mat_wtx, mat_wtw_qinv_lambda))
```

```r
# right hand side
mat_xty <- crossprod(mat_X, vec_y)
mat_wty <- crossprod(mat_W, vec_y)
mat_rhs_marker <- rbind(mat_xty, mat_wty)
```

Solving MME

```r
(mat_sol_marker <- solve(mat_coef_marker, mat_rhs_marker))
```

```
##            [,1]
##      95.7343072
##      92.8583375
## SNP1  3.7968414
## SNP2  8.1559425
## SNP3 -6.8889005
## SNP4  3.5517799
## SNP5  0.4036606
```

Solutions for fixed effects

```r
n_nr_fix_effects <- length(vec_b)
mat_sol_marker[1:n_nr_fix_effects,]
```

```
##
## 95.73431 92.85834
```

The solution for the marker effects

```r
(vec_sol_geno <- mat_sol_marker[(n_nr_fix_effects+1):nrow(mat_sol_marker),])
```

```
##       SNP1       SNP2       SNP3       SNP4       SNP5
##  3.7968414  8.1559425 -6.8889005  3.5517799  0.4036606
```

Predicted genomic breeding values for genotyped animals

```r
mat_W %*% vec_sol_geno
```

```
##           [,1]
## [1,] 12.412505
## [2,] -2.632338
## [3,] -9.822361
## [4,] -6.429180
## [5,] 15.964285
## [6,] 22.853185
```

b) Use a breeding value based model in a single step procedure to predict breeding values for all animals shown in the above dataset. The ratio between residual variance and genomic variance is 11.

Verwenden Sie ein Zuchtwert-basiertes Modell in einem Single-Step Verfahren um genomische Zuchtwerte für alle Tiere im oben gezeigten Datensatz zu schätzen. Das Verhältnis der Restvarianz zur genomischen Varianz beträgt 11.

20

Solution

The genomic breeding values using a breeding value based model are obtained by the following linear mixed effects model

$$y = Xb + Zu + e$$

with y the vector of observations, b the vector of fixed effects, u the vector of random genomic breeding values, e the vector of random residuals and design matrices X and Z. Inserting information from the data leads to

```
# vector y
vec_y <- tbl_p4$P
n_nr_rec_p4 <- length(vec_y)
# matrix X
mat_X <- model.matrix(lm(P ~ 0 + SEX, data = tbl_p4))
attr(mat_X, "assign") <- NULL
attr(mat_X, "contrasts") <- NULL
colnames(mat_X) <- NULL
# vector b
vec_b <- c("b_{f}", "b_{m}")
# martix Z
mat_Z <- diag(1, nrow = n_nr_rec_p4)
# vector g
vec_u <- sapply(1:n_nr_rec_p4, function(x) paste("u_", x, sep = ""))
# vector e
vec_e <- sapply(1:nrow(tbl_p4), function(x) paste("e_", x, sep = ""))
```

$$y = \begin{bmatrix} 115.4 \\ 84.4 \\ 86.8 \\ 76.3 \\ 122.3 \\ 115.8 \end{bmatrix}, X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}, b = \begin{bmatrix} b_f \\ b_m \end{bmatrix}, Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix}, e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix}$$

The expected values are defined as

$$E \begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \end{bmatrix}$$

The variance-covariance matrices

$$var \begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} ZHZ^T + R & ZH & R \\ HZ^T & H & 0 \\ R & 0 & R \end{bmatrix}$$

with

- $var(e) = R = I * \sigma_e^2$,
- $var(u) = H = G * \sigma_u^2$
- σ_e^2 the residual variance component
- σ_u^2 the genomic variance component
- G the genomic relationship matrix

The numerator relationshipmatrix G is

```
tbl_geno <- tbl_p4 %>% select(SNP1, SNP2, SNP3, SNP4, SNP5)
mat_geno <- as.matrix(tbl_geno)
# function to compute genomic relationship matrix
computeMatGrm <- function(pmatData) {
  matData <- pmatData
  # check the coding, if matData is -1, 0, 1 coded, then add 1 to get to 0, 1, 2 coding
  if (min(matData) < 0) matData <- matData + 1
  # Allele frequencies, column vector of P and sum of frequency products
  freq <- apply(matData, 2, mean) / 2
  P <- 2 * (freq - 0.5)
  sumpq <- sum(freq*(1-freq))
  # Changing the coding from (0,1,2) to (-1,0,1) and subtract matrix P
  Z <- matData - 1 - matrix(P, nrow = nrow(matData),
                            ncol = ncol(matData),
                            byrow = TRUE)
  # Z%*%Zt is replaced by tcrossprod(Z)
  return(tcrossprod(Z)/(2*sumpq))
}
# genomic relationship matrix
mat_G <- computeMatGrm(pmatData = mat_geno)
# test full rank
if (Matrix::rankMatrix(mat_G) < nrow(mat_G)){
  mat_G_star <- mat_G + 0.01 * diag(1, nrow = nrow(mat_G))
} else {
  mat_G_star <- mat_G
}
```

$$G = \begin{bmatrix} 0.4286 & -0.0465 & -0.6977 & -0.3256 & 0.2326 & 0.4186 \\ -0.0465 & 0.6147 & -0.0465 & 0.3256 & -0.2326 & -0.6047 \\ -0.6977 & -0.0465 & 2.103 & 0.7907 & -0.8837 & -1.2558 \\ -0.3256 & 0.3256 & 0.7907 & 0.6147 & -0.5116 & -0.8837 \\ 0.2326 & -0.2326 & -0.8837 & -0.5116 & 0.6147 & 0.7907 \\ 0.4186 & -0.6047 & -1.2558 & -0.8837 & 0.7907 & 1.5449 \end{bmatrix}$$

Setting up mixed model equations

```r
# coefficient matrix
mat_xtx <- crossprod(mat_X)
mat_xtz <- crossprod(mat_X, mat_Z)
mat_ztx <- t(mat_xtz)
mat_ztz_ginv_lambda <- crossprod(mat_Z) + lambda_g * solve(mat_G_star)
mat_coef_geno <- rbind(cbind(mat_xtx, mat_xtz), cbind(mat_ztx, mat_ztz_ginv_lambda))
# right hand side
mat_xty <- crossprod(mat_X, vec_y)
mat_zty <- crossprod(mat_Z, vec_y)
mat_rhs_geno <- rbind(mat_xty, mat_zty)
```

Solving MME

```r
(mat_sol_geno <- solve(mat_coef_geno, mat_rhs_geno))
```

```
##            [,1]
## [1,] 101.951174
## [2,]  96.597652
## [3,]   2.410232
## [4,]  -2.155119
## [5,]  -6.107255
## [6,]  -3.905536
## [7,]   3.650593
## [8,]   6.107085
```

Solutions for fixed effects

```r
n_nr_fix_effects <- length(vec_b)
mat_sol_geno[1:n_nr_fix_effects,]
```

```
## [1] 101.95117  96.59765
```

Predicted genomic breeding values for genotyped animals

```r
(vec_sol_geno <- mat_sol_geno[(n_nr_fix_effects+1):nrow(mat_sol_geno),])
```

```
## [1]  2.410232 -2.155119 -6.107255 -3.905536  3.650593  6.107085
```

Problem 5: Contrasts

The dataset shown below is used to fit a fixed linear model with `Methane` as response variable and `Breed` as predictor using different contrasts.

Der nachfolgende Datensatz wird in einem fixen linearen Modell mit 'Methane' als Zielgrösse und 'Breed' als unabhängiger Variablen unter Verwendung verschiedener Kontraste verwendet.

Cow	Methane	Breed
1	24.41	Angus
2	24.03	Angus
3	24.70	Angus
4	23.98	Angus
5	25.34	Angus
6	22.01	Limousin
7	24.83	Limousin
8	23.77	Limousin
9	23.37	Limousin
10	24.33	Limousin
11	24.05	Limousin
12	22.23	Limousin

```
## https://charlotte-ngs.github.io/asmss2023/data/asm_exam_p05.csv
```

a) Fit a fixed linear effects model with 'Methane' as response variable and with 'Breed' as predictor variable. Use the 'treatment' contrast to report the differences in expected levels of 'Methane' for the different breeds. Verify your results by setting up the least squares normal equations and show how 'treatment' contrasts are computed from these solutions.

   Passen Sie ein fixes lineares Modell an die oben gezeigten Daten an mit 'Methane' als Zielgrösse und mit 'Breed' als unabhängiger Variable. Verwenden Sie die 'treatement' Kontraste um die unterschiedlichen erwarteten Methanwerte für die verschiedenen Rassen anzugeben. Verifizieren Sie Ihre Resultate anhand der Least-Squares Normalgleichungen und berechnen Sie die 'treatement' Kontraste aus den Lösungen der Normalgleichungen.

8

Solution

Treatment contrasts are the default setting in `lm()`, hence the fixed linear effects model is fitted as

```
lm_met_br_p5 <- lm(Methane ~ Breed, data = tbl_p5)
smry_lm_met_br_p5 <- summary(lm_met_br_p5)
smry_lm_met_br_p5
```

```
##
## Call:
## lm(formula = Methane ~ Breed, data = tbl_p5)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5029 -0.4745  0.0630  0.6071  1.3171
```

```
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    24.4920     0.3982  61.513 3.13e-14 ***
## BreedLimousin  -0.9791     0.5213  -1.878   0.0898 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8903 on 10 degrees of freedom
## Multiple R-squared:  0.2608, Adjusted R-squared:  0.1869
## F-statistic: 3.528 on 1 and 10 DF,  p-value: 0.08979
```

The least squares normal equations

```
vec_y <- tbl_p5$Methane
mat_X <- model.matrix(lm(Methane ~ 0 + Breed, data = tbl_p5))
attr(mat_X, "assign") <- NULL
attr(mat_X, "contrasts") <- NULL
colnames(mat_X) <- NULL
mat_X <- cbind(matrix(rep(1,nrow(mat_X)), ncol = 1), mat_X)
mat_xtx <- crossprod(mat_X)
mat_xtx_ginv <- MASS::ginv(mat_xtx)
mat_xty <- crossprod(mat_X, vec_y)
mat_b_sol <- crossprod(mat_xtx_ginv, mat_xty)
mat_b_sol
```

```
##           [,1]
## [1,] 16.001619
## [2,]  8.490381
## [3,]  7.511238
```

Contrasts matrix for `treatment` contrasts

```
fac_breed <- as.factor(tbl_p5$Breed)
contr_mat_breed_treat <- contrasts(fac_breed)
contr_mat_breed_treat <- cbind(matrix(rep(1,nrow(contr_mat_breed_treat)), ncol = 1),
                               contr_mat_breed_treat)
est_mat_breed_treat <- solve(contr_mat_breed_treat)
est_mat_breed_treat
```

```
##          Angus Limousin
##              1        0
## Limousin    -1        1
```

The first row of the above matrix `est_mat_breed_treat` shows how the intercept estimate is computed from the observation means. This means that with the treatment contrasts, the intercept is the mean observation for animals of the `Angus` breed. Hence, we get

```
library(dplyr)
tbl_p5_an <- tbl_p5 |>
  filter(Breed == "Angus" )
mean(tbl_p5_an$Methane)
```

```
## [1] 24.492
```

Comparing that to the intercept from `lm()`

```
smry_lm_met_br_p5$coefficients["(Intercept)", "Estimate"]
```

```
## [1] 24.492
```

For the effects estimates, we are looking at the second and the third row of the matrix `est_mat_breed_treat`. We are prepending a column of zeroes to the second and the third row of `est_mat_breed_treat`.

```
n_nrow_est_mat <- nrow(est_mat_breed_treat)
mat_q_efun <- cbind(matrix(0, nrow = (nrow(est_mat_breed_treat)-1), ncol = 1),
                    matrix(est_mat_breed_treat[2:n_nrow_est_mat,],
                           nrow = (nrow(est_mat_breed_treat)-1)))
crossprod(t(mat_q_efun), mat_b_sol)
```

```
##            [,1]
## [1,] -0.9791429
```

These values correspond to the effect estimates from `lm()`

```
n_nr_coef_row <- nrow(smry_lm_met_br_p5$coefficients)
smry_lm_met_br_p5$coefficients[2:n_nr_coef_row, "Estimate"]
```

```
## [1] -0.9791429
```

b) Fit a fixed linear effects model with ‘Methane‘ as response variable and with ‘Breed‘ as predictor variable. Use the ‘Helmert‘ contrast to report the differences in expected levels of ‘Methane‘ for the different breeds. Verify your results by setting up the least squares normal equations and show how ‘Helmert‘ contrasts are computed from these solutions.

Passen Sie ein fixes lineares Modell an die oben gezeigten Daten an mit ‘Methane‘ als Zielgrösse und mit ‘Breed‘ als unabhängiger Variable. Verwenden Sie die ‘Helmert‘ Kontraste um die unterschiedlichen erwarteten Methanwerte für die verschiedenen Rassen anzugeben. Verifizieren Sie Ihre Resultate anhand der Least-Squares Normalgleichungen und berechnen Sie die ‘Helmert‘ Kontraste aus den Lösungen der Normalgleichungen.

8

Solution

Helmert contrasts can be specified as argument in `lm()`, hence the fixed linear effects model is fitted as

```
lm_met_br_p5 <- lm(Methane ~ Breed, data = tbl_p5, contrasts = list(Breed = "contr.helmert"))
smry_lm_met_br_p5 <- summary(lm_met_br_p5)
smry_lm_met_br_p5
```

```
##
## Call:
## lm(formula = Methane ~ Breed, data = tbl_p5, contrasts = list(Breed = "contr.helmert"))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5029 -0.4745  0.0630  0.6071  1.3171
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  24.0024     0.2607  92.084 5.58e-16 ***
## Breed1       -0.4896     0.2607  -1.878   0.0898 .
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 0.8903 on 10 degrees of freedom
## Multiple R-squared:  0.2608, Adjusted R-squared:  0.1869
## F-statistic: 3.528 on 1 and 10 DF,  p-value: 0.08979
```

The least squares normal equations

```
vec_y <- tbl_p5$Methane
mat_X <- model.matrix(lm(Methane ~ 0 + Breed, data = tbl_p5))
attr(mat_X, "assign") <- NULL
attr(mat_X, "contrasts") <- NULL
colnames(mat_X) <- NULL
mat_X <- cbind(matrix(rep(1,nrow(mat_X)), ncol = 1), mat_X)
mat_xtx <- crossprod(mat_X)
mat_xtx_ginv <- MASS::ginv(mat_xtx)
mat_xty <- crossprod(mat_X, vec_y)
mat_b_sol <- crossprod(mat_xtx_ginv, mat_xty)
mat_b_sol
```

```
##           [,1]
## [1,] 16.001619
## [2,]  8.490381
## [3,]  7.511238
```

Contrasts matrix for `treatment` contrasts

```
fac_breed <- as.factor(tbl_p5$Breed)
contr_mat_breed_helm <- contrasts(C(fac_breed, helmert))
contr_mat_breed_helm <- cbind(matrix(rep(1,nrow(contr_mat_breed_helm)), ncol = 1),
                              contr_mat_breed_helm)
est_mat_breed_helm <- solve(contr_mat_breed_helm)
est_mat_breed_helm
```

```
##      Angus Limousin
## [1,]   0.5      0.5
## [2,]  -0.5      0.5
```

The first row of the above matrix `est_mat_breed_helm` shows how the intercept estimate is computed from the observation means. This means that with the Helmert contrasts, the intercept is the weighted mean of the mean observation for animals of the different breeds. Hence, we get

```
library(dplyr)
tbl_p5_an <- tbl_p5 |>
  filter(Breed == "Angus" )
tbl_p5_li <- tbl_p5 |>
  filter(Breed == "Limousin")
n_nr_breed <- nlevels(as.factor(tbl_p5$Breed))
(mean(tbl_p5_an$Methane) + mean(tbl_p5_li$Methane)) / n_nr_breed
```

```
## [1] 24.00243
```

Comparing that to the intercept from `lm()`

```
smry_lm_met_br_p5$coefficients["(Intercept)", "Estimate"]
```

```
## [1] 24.00243
```

For the effects estimates, we are looking at the second and the third row of the matrix `est_mat_breed_helm`. We are prepending a column of zeroes to the second and the third row of `est_mat_breed_helm`.

```
n_nrow_est_mat <- nrow(est_mat_breed_helm)
mat_q_efun <- cbind(matrix(0, nrow = (nrow(est_mat_breed_helm)-1), ncol = 1),
                    matrix(est_mat_breed_helm[2:n_nrow_est_mat,],
                           nrow = (nrow(est_mat_breed_helm)-1)))
crossprod(t(mat_q_efun), mat_b_sol)
```

```
##            [,1]
## [1,] -0.4895714
```

These values correspond to the effect estimates from `lm()`

```r
n_nr_coef_row <- nrow(smry_lm_met_br_p5$coefficients)
smry_lm_met_br_p5$coefficients[2:n_nr_coef_row, "Estimate"]
```

```
## [1] -0.4895714
```