

Züchtungslehre – Übung 9

Lösung

11. Dezember 2015

Aufgabe

Es soll eine genomische Zuchtwertschätzung mit der Methode SNP-BLUP durchgeführt werden. Die Referenzpopulation besteht aus 325 Stieren, welche sicher geschätzte Zuchtwerte für das Merkmal Eiweiss-% aufweisen. Die Stiere wurden für 10 SNP genotypisiert. Die Genotypen befinden sich in der Designmatrix in der Datei **xmatrix.txt**, die Phänotypen in der Datei **yvect.txt**.

Bei 31 Stierkälbern stehen Selektionsentscheide für den Testeinsatz an. Die Kälber sind ebenfalls für die gleichen 10 SNP Marker genotypisiert. Die Genotypen der Kälber sind in der Designmatrix in der Datei **x_kalb.dat**.

a) Für die 31 Kälber sollen direkt genomische Zuchtwerte für das Merkmal Eiweiss-% geschätzt werden. Die SNP-Effekte für die 10 SNP werden nach folgenden Mischmodellgleichungen simultan geschätzt.

$$\begin{bmatrix} \mathbf{1}_n' \mathbf{1}_n & \mathbf{1}_n' \mathbf{X} \\ \mathbf{X}' \mathbf{1}_n & \mathbf{X}' \mathbf{X} + \mathbf{I}\lambda \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n' \mathbf{y} \\ \mathbf{X}' \mathbf{y} \end{bmatrix}$$

Es wird angenommen, dass der Varianzenquotient $\lambda = 10$ ist.

Der direkt genomische Zuchtwert der Kälber errechnet sich aus

$$DGZW = X\hat{\mathbf{g}}$$

Lösen Sie das Gleichungssystem. Wie ist die Lösung für den Mittelwert? Welcher SNP hat den grössten Effekt? Berechnen Sie die direkt genomischen Zuchtwerte für die Kälber.

Anleitung:

Legen Sie zuerst einige Parameter fest: `nmarkers<-10;nrecords<-325; lamda<-10`

Lesen Sie anschliessend die Dateien als Matrix ein.

Der Vektor mit „1“ für den Mittelwert und die Einheitsmatrix können folgendermassen angelegt werden: `ones<-array(1,c(nrecords)); ident_mat<-diag(nmarkers)`

Im nächsten Schritt muss die Koeffizientenmatrix gebildet werden. Dies kann in Blöcken gemacht werden. Zuerst soll die Dimension der Koeffizientenmatrix festgelegt werden: `coeff<-array(0,c(nmarkers+1,nmarkers+1))`

Diese kann nun befüllt werden:

```
Coeff[1:1,1:1]<-t(ones)%*%ones
```

```
Coeff[1:1,2:nmarkers+1]<-t(ones)%*%X (X ist die eingelesen Designmatrix mit den Genotypen)
```

Ähnlich dieser Vorgangsweise sollen nun die weiteren Blöcke der Koeffizientenmatrix sowie $1'y$ und $X'y$ erstellt werden.

b) Vier Jahre später erhalten die Kälber einen Zuchtwert aus der Nachkommenprüfung (traditioneller Zuchtwert). Die Zuchtwerte befinden sich in der Datei **y_kalb.dat**. Vergleichen Sie die genomischen Zuchtwerte mit den traditionellen Zuchtwerten und berechnen Sie die Korrelation zwischen beiden.

Lösung

#Pfad setzen:

```
setwd("/Pfad_zu_den_Daten/")
```

#Parameter setzen

```
nmarkers<-10
```

```
nrecords<-325
```

```
lamda<-10
```

#Daten einlesen

```
x<-matrix(scan("xmatrix.txt"),ncol=nmarkers,byrow=T)
```

```
y<-matrix(scan("yvec.txt"), byrow=T)
```

```
xkalb <- matrix(scan("x_kalb.dat"),ncol=nmarkers, byrow=T)
```

```
ykalb <- matrix(scan("y_kalb.dat"), byrow=T)
```

#Vektor mit „1“

```
ones<-array(1,c(nrecords))
```

#Einheitsmatrix

```
ident_matrix <- diag(nmarkers)
```

#Dimension der Koeffizientenmatrix festlegen

```
coeff <- array(0,c(nmarkers+1, nmarkers+1))
```

#Koeffizientenmatrix mit Submatrizen befüllen

```
coeff[1:1,1:1]<- t(ones)%*%ones
```

```
coeff[1:1,2:(nmarkers+1)] <- t(ones)%*%x
```

```
coeff[2:(nmarkers+1),1:1] <- t(x)%*%ones
```

```
xx <- t(x)%*%x
```

```
i10 <- ident_matrix*lamda
```

```
xx1 <- xx + i10
```

```
coeff[2:(nmarkers+1),2:(nmarkers+1)] <- xx1
```

```
#1'1
```

```
#1'X
```

```
#X'1
```

```
#X'X
```

```
#I*λ
```

```
#X'X+I*λ
```

#Dimension Right hand side festlegen

```
rhs <- array(0,(nmarkers+1))
```

#Right hand side befüllen

```
rhs[1] <- t(ones)%**y
```

```
#1'y
```

```
rhs[2:(nmarkers+1)] <- t(x)%**y
```

```
#X'y
```

#Gleichungssystem lösen

```
solution_vec <- solve(coeff, rhs)
```

#SNP-Effekte (g-hat) selektieren

```
g_hat <- solution_vec[2:(nmarkers+1)]
```

#DGZW für die 31 Kälber berechnen

```
dgzw <- xkalb%**g_hat
```

#Korrelation zwischen DGZW und Zuchtwert aus der Nachkommenprüfung (nach 4 Jahren)

```
cor(dgzw, ykalb)
```