3 Generalisierte lineare Modelle

Lösung zu Aufgabe 4

(a) Matrix $D(\boldsymbol{\beta})$ für das Logit Modell:

$$D(\boldsymbol{\beta}) = diag\left(\frac{\partial h(\eta_1)}{\partial \eta}, \dots, \frac{\partial h(\eta_n)}{\partial \eta}\right),$$

wobei

$$\frac{\partial h(\eta_j)}{\partial \eta} = \frac{\partial \frac{exp(\eta_j)}{1+exp(\eta_j)}}{\partial \eta} = \frac{exp(\eta_j)}{(1+exp(\eta_j))^2}.$$

Im Logit-Modell gilt außerdem:

$$W(\boldsymbol{\beta}) = D(\boldsymbol{\beta})$$

```
# Funktion, die IWLS fürs Logit-Modell durchführt
iwls_logit <- function(y,X,beta0,eps=1e-5,maxit=100){

  #Argumente:
  #y        = Response-Vektor, (n x 1)
  #X        = Matrix mit Kovariablen, (n x p)
  #beta0    = Startwert für den Parametervektor
  #eps      = Toleranzgrenze
  #maxit    = maximale Anzahl der Iterationen

  #Initialisierung:
  X <- as.matrix(X)
  betaOld <- beta0          # Parametervektor
  i <- 0                    # Zähler
  conv <- FALSE             # Indikator für Konvergenz

  while(conv==FALSE){ #Schleife

    #Initialisierung/update
    eta <- X%*%betaOld                    # linearer Prädiktor
    mu <- exp(eta)/(1+exp(eta))           # Vektor der gefitteten Werte
    d  <- exp(eta)/(1+exp(eta))^2         # Ableitung von mu
    W  <- diag(as.vector(d))              # Gewichtsmatrix
    pseudo <- eta+(y-mu)/d                # PseudoBeobachtungen
    Finv <- solve(t(X)%*%W%*%X)           # Inverse Fisher-Matrix

    #IWLS in k-ter Iteration
    betaNew <- Finv%*%t(X)%*%W%*%pseudo
    i <- i+1

    if(sqrt(sum((betaNew-betaOld)^2)/sum(betaOld^2)) < eps | i>=maxit){
      #Überprüfung des Abbruchkriteriums
      conv <- TRUE
    }

    betaOld <- betaNew
  }
```

```
  #update zur Berechnung von eta, mu und cov im terminalen Wert
  eta <- X%*%betaNew
  mu <- exp(eta)/(1+exp(eta))
  d  <- exp(eta)/(1+exp(eta))^2
  W  <- diag(as.vector(d))
  Finv <- solve(t(X)%*%W%*%X)

  #Rückgabe
  result <- list("coef"=betaNew,       # geschätzter Parametervektor
                 "fitted"=mu,         # gefittete Werte
                 "predictor"=eta,     # geschätzter lin. Prädiktor
                 "cov"=Finv,          # Kovarianzmatrix von \hat{\beta}
                 "no.its"=i)          # Anzahl Iterationen
  return(result)
}
```

(b) Analyse des Datensatzes shuttle von Übungsblatt 2

```
shuttle <- read.table("shuttle.asc", header = TRUE)

# Designmatrix und Response
y <- shuttle$td
X <- cbind(1,shuttle$temp)

# mit Startwert (0,0)
mod1 <- iwls_logit(y,X,beta0=c(0,0))
mod1$coef

##             [,1]
## [1,] 15.0429016
## [2,] -0.2321627

mod1$no.its # 6 Iterationen

## [1] 6

# mit Startwert (4,-0.1)
mod2 <- iwls_logit(y,X,beta0=c(4,-0.1))
mod2$coef

##             [,1]
## [1,] 15.0429016
## [2,] -0.2321627

mod2$no.its # 7 Iterationen

## [1] 7

# mit Startwert (2,-2)
mod3 <- iwls_logit(y,X,beta0=c(2,-2))

## Error in if (sqrt(sum((betaNew - betaOld)^2)/sum(betaOld^2)) < eps | i >= : Fehlender
Wert, wo TRUE/FALSE nötig ist
```

Mit Startwert c(4,-0.1) benötigt der Algorithmus eine Iteration mehr, obwohl die Startwerte näher am ML-Schätzer liegen. Mit den Startwerten c(2,-2) konvergiert der Algorithmus nicht. Man erkennt, die starke Abhängigkeit des Algorithmus von Startwerten.

```r
# Modelle mit glm()
mod_glm1 <- glm(td ~ temp, data = shuttle, family=binomial)
mod_glm1$coefficients
```

```
## (Intercept)        temp
##  15.0429016  -0.2321627
```

```r
mod_glm2 <- glm(td ~ temp, data = shuttle, family=binomial, start=c(2,-2))
summary(mod_glm2)
```

```
##
## Call:
## glm(formula = td ~ temp, family = binomial, data = shuttle, start = c(2,
##     -2))
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
##   0.00    0.00    0.00    0.00    8.49
##
## Coefficients:
##               Estimate Std. Error   z value Pr(>|z|)
## (Intercept)  8.221e+15  1.417e+08  58003512   <2e-16 ***
## temp        -1.322e+14  2.027e+06 -65230006   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance:  28.267  on 22  degrees of freedom
## Residual deviance: 288.349  on 21  degrees of freedom
## AIC: 292.35
##
## Number of Fisher Scoring iterations: 25
```

Mit den Startwerten c(2,-2) konvergiert auch die eingebaute Schätzfunktion nicht.