Project1

Erle E. Sandø, Simon Liabø

2022-09-23

Problem 1

1

a)

b)

```
score = function(y, X, beta)
  {
  eta = as.vector(X %*% beta)
  lmdba = exp(eta)
  score = apply((y - lmdba) * X,2,sum)
  score
}
expected_fisher = function(X, beta)
  {
  eta = as.vector(X %*% beta)
  W = diag(exp(eta))
  t(X) %*% W %*% X
}
log_likelihood = function(y, X, beta, lmbda = exp(as.vector(X %*% beta)))
  {
  sum(ifelse(lmbda==0,0,y*log(lmbda) ) - lmbda)
}
myglm = function(formula, data, start=rep(0, ncol(model.matrix(formula, data))))
  {
  X = model.matrix(formula, data)
  response = as.character(formula)[2]
  y = data[[response]]
  beta = start

  s=1
  while (s > (1e-10)) {
    eta = as.vector(X %*% beta)
    lmbda = exp(eta)

    score_val = score(y, X, beta)
    f = expected_fisher(X, beta)

    beta = beta + solve(f) %*% score_val
    s = sum(score_val^2)
    }
```

```
  #vcov
  cov_mat = solve(f)

  #coefficients
  sd_err = sqrt(diag(cov_mat))
  coeff = cbind(beta, sd_err)
  colnames(coeff) = c("Estimate", "Std.Error")
  rownames(coeff) = paste0("beta_", seq_along(beta)-1)

  #deviance
  dev = 2 * (log_likelihood(y, X, beta) - log_likelihood(y, X, beta, lmbda = y))

  list(coefficients = coeff, deviance = dev, vcov = cov_mat)
}
```

c)

```
n = 1000
k = 2

#simulate data
beta = rnorm(k+1)
X = cbind(matrix(1,n),matrix(rnorm(n * k), nrow = n, ncol = k))
eta = as.vector(X %*% beta)

lmd = exp(eta)
y = rpois(n,lmd)

data_sim = as.data.frame(cbind(y,X[,2:3]))

#fit models
model_myglm = myglm(y~., data = data_sim)
model_glm = glm(y~., data = data_sim, family = poisson(link=log))

#evaluate
coeff_diff = mean((model_myglm$coefficients[,1] -model_glm$coefficients)^2)
coeff_diff
```

```
## [1] 5.693269e-21
```

```
vcov_diff = mean( (model_myglm$vcov - vcov(model_glm))^2)
vcov_diff
```

```
## [1] 9.152911e-17
```

The model looks good. The results are very close to those obtained with glm() and vcov().

d)

```
load(url("https://www.math.ntnu.no/emner/TMA4315/2022h/hoge-veluwe.Rdata"))
model_myglm = myglm(y ~ t + I(t^2), data)
model_myglm
```

```
## $coefficients
##             Estimate   Std.Error
```

```
## beta_0  1.420130461 0.282434733
## beta_1  0.085183057 0.034053955
## beta_2 -0.003298608 0.001019464
##
## $deviance
## [1] -277.4613
##
## $vcov
##               (Intercept)             t        I(t^2)
## (Intercept)  0.0797693783 -9.308596e-03  2.550195e-04
## t           -0.0093085957  1.159672e-03 -3.369024e-05
## I(t^2)       0.0002550195 -3.369024e-05  1.039306e-06
```