

# Issue-based Congressional Network Regression Analysis

Lynette Dang

2022-06-01

```
library(readr)
df <- read_csv("issue_analysis.csv")

## New names:
## Rows: 481 Columns: 11
## -- Column specification
## ----- Delimiter: "," chr
## (5): Legislator name, Gender Estimates, Predicted Gender, Ethnicity Esti... dbl
## (6): ...1, Legislative Effectiveness Score, bc, ec, cc, dc
## i Use 'spec()' to retrieve the full column specification for this data. i
## Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## * ' -> '...1'

df[is.na(df)] = 0
df$factor <- as.factor(sample(1:2, nrow(df), replace = TRUE))
head(df)

## # A tibble: 6 x 12
##   ...1 'Legislator name' 'Legislative Effe~' 'Gender Estima~' 'Predicted Gen~'
##   <dbl> <chr>           <dbl> <chr>           <chr>
## 1     0 Jack Reed      0.346 {'male': 0.1144~ male
## 2     1 Kirsten Gillibrand 0.339 {'male': 0.0001~ female
## 3     2 Tammy Duckworth  1.53  {'male': 0.0004~ male
## 4     3 Gary Peters      5.02  {'male': 0.0014~ male
## 5     4 Elizabeth Warren  0.725 {'male': 8.0722~ female
## 6     5 Martin Heinrich  0.781 {'male': 0.8787~ male
## # ... with 7 more variables: 'Ethnicity Estimates' <chr>,
## #   'Predicted Ethnicity' <chr>, bc <dbl>, ec <dbl>, cc <dbl>, dc <dbl>,
## #   factor <fct>

library(glmnet)

## Loading required package: Matrix
## Loaded glmnet 4.1-3

library(xtable)

# the original data frame and formula
set.seed(23)
```

```
f <- as.formula(df$`Legislative Effectiveness Score` ~ (df$bc+ df$ec+ df$cc+ df$dc + df$`Predicted Gender`))
# transform dataframe to matrices as required by glmnet
x <- model.matrix(f, df)
y <- as.matrix(df$`Legislative Effectiveness Score`, ncol=1)

# fit ridge regression model with a wide range of penalty parameter
lambdas <- 10^seq(3, -2, by = -.1)
cv_fit <- cv.glmnet(x,y, lambda=lambdas, alpha = 0)
fit <- cv_fit$glmnet.fit
summary(fit)
```

```
##           Length Class      Mode
## a0           51   -none-   numeric
## beta        1734 dgCMatrix S4
## df           51   -none-   numeric
## dim           2   -none-   numeric
## lambda       51   -none-   numeric
## dev.ratio    51   -none-   numeric
## nulldev       1   -none-   numeric
## npasses       1   -none-   numeric
## jerr          1   -none-   numeric
## offset        1   -none-   logical
## call          5   -none-   call
## nobs          1   -none-   numeric
```

```
# get BIC value
tLL <- fit$nulldev - deviance(fit)
k <- fit$df
n <- fit$nobs

BIC<-log(n)*k - tLL
BIC
```

```
## [1] 203.7616 203.7509 203.7373 203.7203 203.6990 203.6723 203.6390 203.5974
## [9] 203.5457 203.4814 203.4018 203.3038 203.1836 203.0372 202.8641 202.6532
## [17] 202.4026 202.1080 201.7656 201.3721 200.9253 200.4236 199.8659 199.2522
## [25] 198.5828 197.8586 197.0818 196.2561 195.3874 194.4868 193.5567 192.6122
## [33] 191.6583 190.7047 189.7591 188.8233 187.8986 186.9835 186.0519 185.1488
## [41] 184.1809 183.2578 182.2134 181.2460 180.1115 179.1123 177.9033 176.9283
## [49] 175.8363 174.6532 173.8858
```

```
opt_lambda = lambdas[which(BIC == min(BIC))]
# Choose best model
best_model <- glmnet(x,y, lambda=opt_lambda, alpha = 0)
df <- data.frame(coef.name = dimnames(coef(best_model))[[1]], coef.value = matrix(coef(best_model)))
xtable(df)
```

```
## % latex table generated in R 4.1.2 by xtable 1.8-4 package
## % Mon May 30 21:51:07 2022
## \begin{table}[ht]
## \centering
## \begin{tabular}{rlr}
```

```

## \hline
## & coef.name & coef.value \\
## \hline
## 1 & (Intercept) & 1.16 \\
## 2 & (Intercept) & 0.00 \\
## 3 & df\${bc} & -145.13 \\
## 4 & df\${ec} & 5.06 \\
## 5 & df\${cc} & 1.15 \\
## 6 & df\${dc} & -0.16 \\
## 7 & df\${'Predicted Gender'male} & 0.21 \\
## 8 & df\${'Predicted Ethnicity'black} & -0.66 \\
## 9 & df\${'Predicted Ethnicity'hispanic} & -0.03 \\
## 10 & df\${'Predicted Ethnicity'white} & -0.48 \\
## 11 & df\${bc:df\${ec}} & 1859.87 \\
## 12 & df\${bc:df\${cc}} & 172.04 \\
## 13 & df\${bc:df\${dc}} & 148.41 \\
## 14 & df\${bc:df\${'Predicted Gender'male}} & 21.59 \\
## 15 & df\${bc:df\${'Predicted Ethnicity'black}} & 411.98 \\
## 16 & df\${bc:df\${'Predicted Ethnicity'hispanic}} & 147.50 \\
## 17 & df\${bc:df\${'Predicted Ethnicity'white}} & 42.75 \\
## 18 & df\${ec:df\${cc}} & 7.00 \\
## 19 & df\${ec:df\${dc}} & 3.88 \\
## 20 & df\${ec:df\${'Predicted Gender'male}} & 2.03 \\
## 21 & df\${ec:df\${'Predicted Ethnicity'black}} & -1558.70 \\
## 22 & df\${ec:df\${'Predicted Ethnicity'hispanic}} & -557.74 \\
## 23 & df\${ec:df\${'Predicted Ethnicity'white}} & -19.37 \\
## 24 & df\${cc:df\${dc}} & -1.59 \\
## 25 & df\${cc:df\${'Predicted Gender'male}} & 0.12 \\
## 26 & df\${cc:df\${'Predicted Ethnicity'black}} & -3.62 \\
## 27 & df\${cc:df\${'Predicted Ethnicity'hispanic}} & 0.04 \\
## 28 & df\${cc:df\${'Predicted Ethnicity'white}} & -0.08 \\
## 29 & df\${dc:df\${'Predicted Gender'male}} & -0.61 \\
## 30 & df\${dc:df\${'Predicted Ethnicity'black}} & 1.95 \\
## 31 & df\${dc:df\${'Predicted Ethnicity'hispanic}} & 0.06 \\
## 32 & df\${dc:df\${'Predicted Ethnicity'white}} & 0.28 \\
## 33 & df\${'Predicted Gender'male:df\${'Predicted Ethnicity'black}} & -0.05 \\
## 34 & df\${'Predicted Gender'male:df\${'Predicted Ethnicity'hispanic}} & -0.15 \\
## 35 & df\${'Predicted Gender'male:df\${'Predicted Ethnicity'white}} & 0.14 \\
## \hline
## \end{tabular}
## \end{table}

```