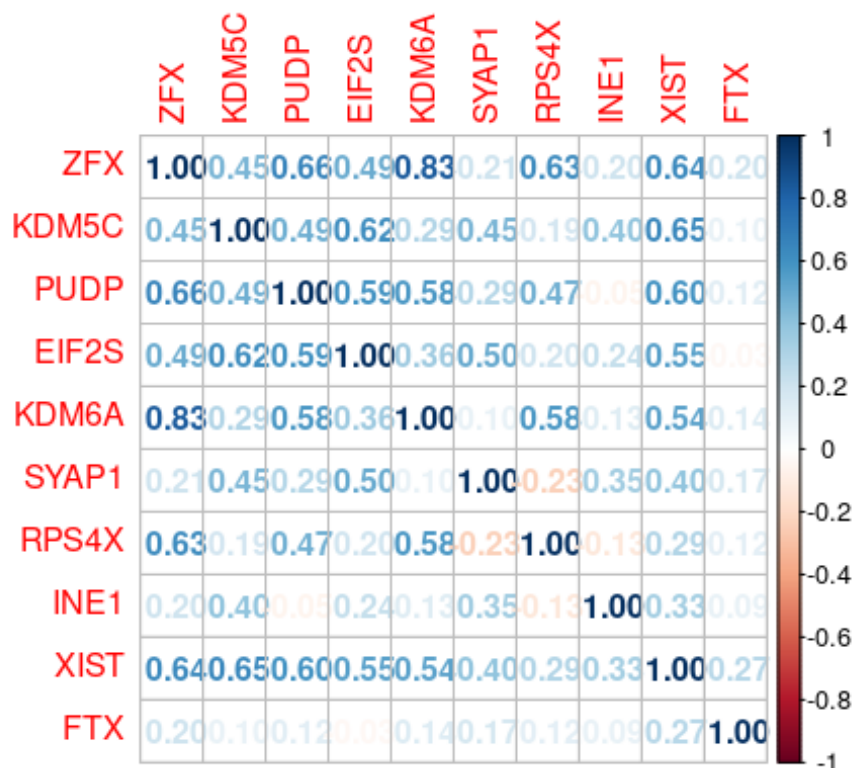


trs_demo2.R

alexander

2023-02-07

```
library(glmnet)
## Loading required package: Matrix
## Loaded glmnet 4.1-6
library(glmnetUtils)
##
## Attaching package: 'glmnetUtils'
## The following objects are masked from 'package:glmnet':
##
##      cv.glmnet, glmnet
library(corrplot)
## corrplot 0.92 loaded
corrplot(M2, method = 'number')
```



Glmnet

```
hits_cov_y_rosmay_mayo.data <- cbind(hits_cov_rosmay_mayo.data,  
wgs_info_transcriptome_rosmay_mayo.data[, "FK_final_CaseControl",  
drop=F])
```

```
# Regress AD status on 10 transcripts, sex, age and cohort
```

```
# Regular logistic regression
```

```
summary(glm(FK_final_CaseControl ~ ., data =  
hits_cov_y_rosmay_mayo.data, family = binomial))
```

```
##
```

```
## Call:
```

```
## glm(formula = FK_final_CaseControl ~ ., family = binomial, data =  
hits_cov_y_rosmay_mayo.data)
```

```
##
```

```
## Deviance Residuals:
```

```
##      Min       1Q   Median       3Q      Max  
## -2.0032  -1.2130   0.7664   0.9359   1.8817
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)  
## (Intercept)   -5.81389    1.68572  -3.449 0.000563 ***  
## ENSG00000005889.15  0.40552    0.30311   1.338 0.180949  
## ENSG00000126012.11  0.34582    0.25450   1.359 0.174205  
## ENSG00000130021.13 -0.01209    0.22210  -0.054 0.956595  
## ENSG00000130741.10  0.16218    0.23090   0.702 0.482439  
## ENSG00000147050.14  0.09469    0.26700   0.355 0.722847  
## ENSG00000169895.5  -0.30205    0.19608  -1.540 0.123446  
## ENSG00000198034.10  0.08623    0.22615   0.381 0.702963  
## ENSG00000224975.1  -0.01730    0.17101  -0.101 0.919433  
## ENSG00000229807.9  -0.23823    0.21439  -1.111 0.266472  
## ENSG00000270223.1   0.03142    0.14800   0.212 0.831874  
## gender          -0.30317    0.55999  -0.541 0.588245  
## age              0.07672    0.01918   4.001 6.31e-05 ***  
## mayo            -0.18396    0.24970  -0.737 0.461279
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
```

```
##      Null deviance: 490.28  on 363  degrees of freedom
```

```
## Residual deviance: 451.58  on 350  degrees of freedom
```

```
## AIC: 479.58
```

```
##
```

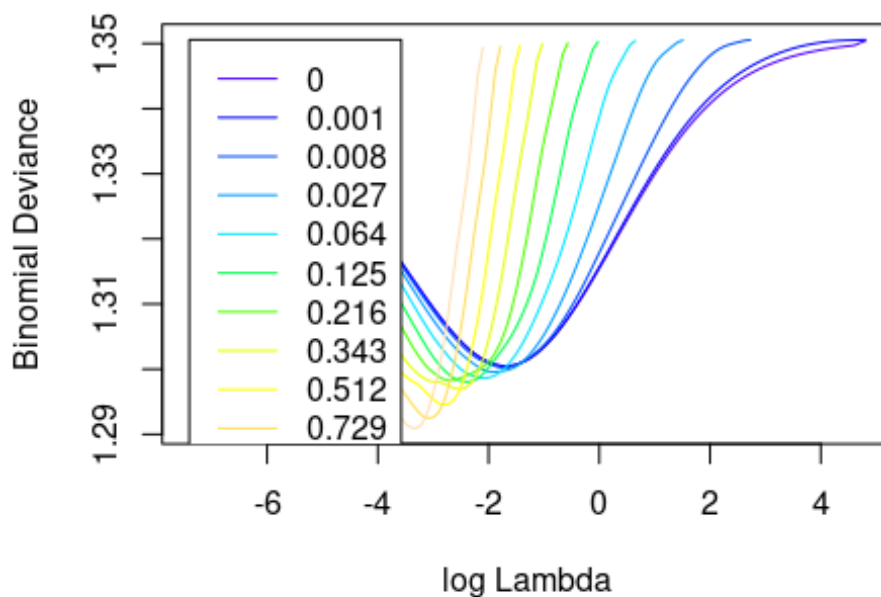
```
## Number of Fisher Scoring iterations: 4
```

```

# Equivalent elastic net
# 10-Fold cross-validation to determine optimal combination of alpha
and lambda
set.seed(20200516)
rosmay_mayo.fit <- cva.glmnet(FK_final_CaseControl ~ ., data =
hits_cov_y_rosmay_mayo.data, family = "binomial")

# Examine which combination of alpha and lambda results in lowest
deviance
plot(rosmay_mayo.fit)

```



```

# Get optimal alpha
# Function from https://stackoverflow.com/a/62154472
get_alpha <- function(fit) {
  alpha <- fit$alpha
  error <- sapply(fit$modlist, function(mod) {min(mod$cvm)})
  alpha[which.min(error)]
}

get_alpha(rosmay_mayo.fit)
## [1] 1

```

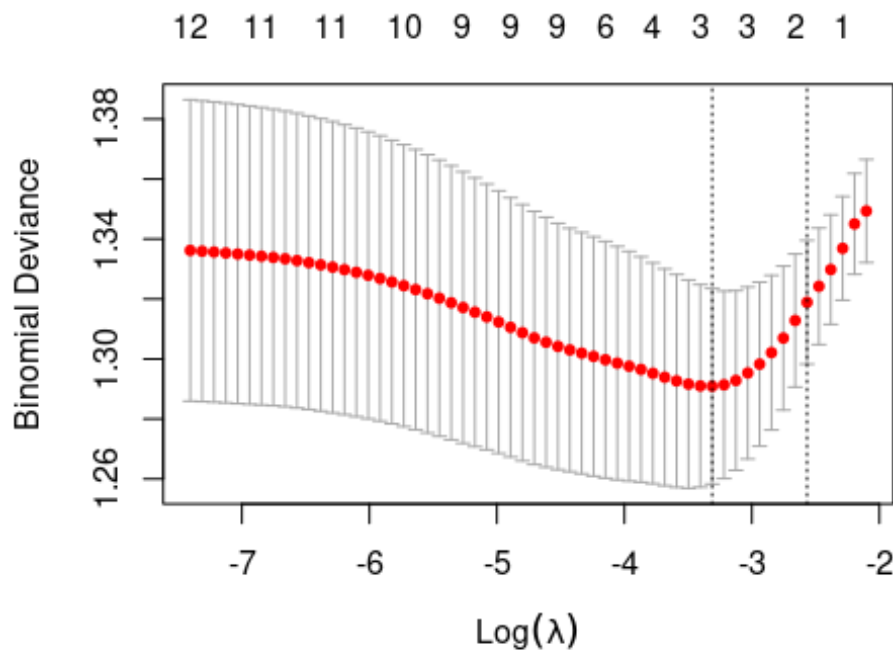
```

# Coefficients at optimal alpha and lambda
coef(rosmay_fit, alpha = 1, s = "lambda.min")

## 14 x 1 sparse Matrix of class "dgCMatrix"
##               s1
## (Intercept)    -3.91948086
## ENSG0000005889.15  0.15671333
## ENSG00000126012.11 .
## ENSG00000130021.13 .
## ENSG00000130741.10 .
## ENSG00000147050.14 .
## ENSG00000169895.5 .
## ENSG00000198034.10  0.07803799
## ENSG00000224975.1 .
## ENSG00000229807.9 .
## ENSG00000270223.1 .
## gender          .
## age              0.05100053
## mayo            .

# Examine different lambda options at optimal alpha (1)
plot(rosmay_fit$modlist[[1]])

```



```
# Examine the shrinking of coefficients
best_apha.fit <- glmnet(hits_cov_rosmay_mayo.data,
wgs_info_transcriptome_rosmay_mayo.data$FK_final_CaseControl, family =
"binomial")
plot(best_apha.fit)
```

