Regularization

Joshua Eubanks

2022-07-25

Contents


```
library(gamlr)
```

```
## Loading required package: Matrix
```

```
sessionInfo()
```

```
## R version 4.2.1 (2022-06-23 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19044)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United States.utf8
## [2] LC_CTYPE=English_United States.utf8
## [3] LC_MONETARY=English_United States.utf8
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.utf8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] gamlr_1.13-7 Matrix_1.4-1
##
## loaded via a namespace (and not attached):
##  [1] lattice_0.20-45 digest_0.6.29   grid_4.2.1      magrittr_2.0.3
##  [5] evaluate_0.15   rlang_1.0.3     stringi_1.7.8   cli_3.3.0
##  [9] rmarkdown_2.14  tools_4.2.1     stringr_1.4.0   xfun_0.31
## [13] yaml_2.3.5      fastmap_1.1.0   compiler_4.2.1  htmltools_0.5.2
## [17] knitr_1.39
```

Regularization is essentially another way to say desensitization. What we will do is trade off bias (deviations from minimizing SSR) and variance (variation in predicted values).

Examples

We are going to go through multiple examples.

NHL

The data comprise of information about play configuration and the players on ice (including goalies) for every goal from 2002-03 to 2012-14 NHL seasons. Collected using A. C. Thomas's `nlhscrapr` package. See the Chicago hockey analytics project at `github.com/mataddy/hockey`

- goal -> Info about each goal scored, including homegoal – an indicator for the home team scoring.
- player -> Sparse Matrix with entries for who was on the ice for each goal: +1 for a home team player, -1 for an away team player, zero otherwise.
- team -> Sparse Matrix with indicators for each team*season interaction: +1 for home team, -1 for away team.
- config -> Special teams info. For example, S5v4 is a 5 on 4 powerplay, +1 if it is for the home-team and -1 for the away team

Data Cleaning

Bringing in data

```
data(hockey)
```

```
x <- cbind(config,team,player) # cbind binds together two sparse matrices
```

Generate Independent Variables

```
y <- goal$homegoal
```

Generating dependent variable, Binary response

Generate Model

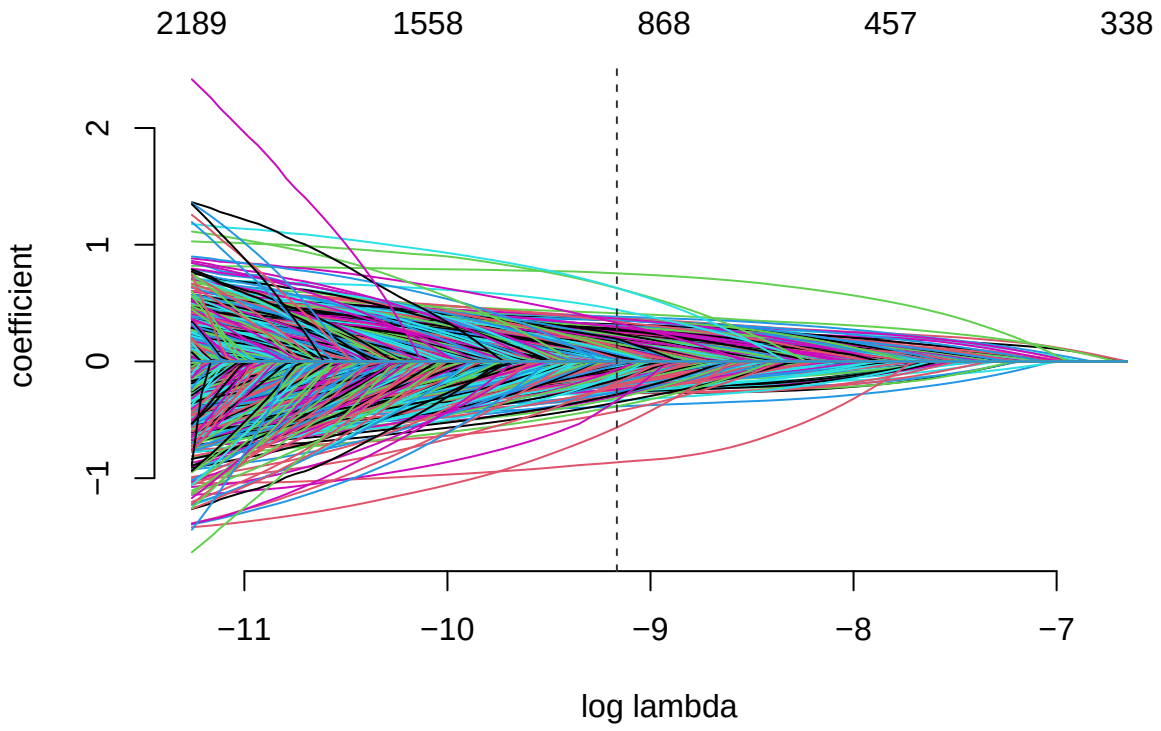
```
nhlreg <- gamlr(x, y, verb=TRUE,
    free=1:(ncol(config)+ncol(team)), ## free denotes unpenalized columns
    family="binomial", standardize=FALSE)
```

```
## *** n=69449 observations and p=2776 covariates ***
## segment 1: lambda = 0.001289, dev = 8.096e+04, npass = 292
## segment 2: lambda = 0.001231, dev = 8.095e+04, npass = 3
## segment 3: lambda = 0.001175, dev = 8.094e+04, npass = 3
## segment 4: lambda = 0.001121, dev = 8.094e+04, npass = 3
## segment 5: lambda = 0.00107, dev = 8.093e+04, npass = 4
## segment 6: lambda = 0.001022, dev = 8.092e+04, npass = 4
## segment 7: lambda = 0.0009753, dev = 8.091e+04, npass = 3
## segment 8: lambda = 0.000931, dev = 8.089e+04, npass = 3
## segment 9: lambda = 0.0008887, dev = 8.087e+04, npass = 4
## segment 10: lambda = 0.0008483, dev = 8.084e+04, npass = 4
## segment 11: lambda = 0.0008097, dev = 8.081e+04, npass = 4
## segment 12: lambda = 0.0007729, dev = 8.079e+04, npass = 4
## segment 13: lambda = 0.0007378, dev = 8.076e+04, npass = 4
## segment 14: lambda = 0.0007043, dev = 8.073e+04, npass = 4
## segment 15: lambda = 0.0006723, dev = 8.07e+04, npass = 4
## segment 16: lambda = 0.0006417, dev = 8.066e+04, npass = 4
## segment 17: lambda = 0.0006125, dev = 8.063e+04, npass = 4
## segment 18: lambda = 0.0005847, dev = 8.059e+04, npass = 4
## segment 19: lambda = 0.0005581, dev = 8.055e+04, npass = 4
```

```
## segment 20: lambda = 0.0005327, dev = 8.051e+04, npass = 4
## segment 21: lambda = 0.0005085, dev = 8.046e+04, npass = 4
## segment 22: lambda = 0.0004854, dev = 8.041e+04, npass = 4
## segment 23: lambda = 0.0004634, dev = 8.036e+04, npass = 4
## segment 24: lambda = 0.0004423, dev = 8.031e+04, npass = 4
## segment 25: lambda = 0.0004222, dev = 8.025e+04, npass = 4
## segment 26: lambda = 0.000403, dev = 8.019e+04, npass = 4
## segment 27: lambda = 0.0003847, dev = 8.013e+04, npass = 4
## segment 28: lambda = 0.0003672, dev = 8.007e+04, npass = 4
## segment 29: lambda = 0.0003505, dev = 8.001e+04, npass = 4
## segment 30: lambda = 0.0003346, dev = 7.995e+04, npass = 4
## segment 31: lambda = 0.0003194, dev = 7.99e+04, npass = 4
## segment 32: lambda = 0.0003049, dev = 7.984e+04, npass = 4
## segment 33: lambda = 0.000291, dev = 7.978e+04, npass = 4
## segment 34: lambda = 0.0002778, dev = 7.972e+04, npass = 4
## segment 35: lambda = 0.0002651, dev = 7.967e+04, npass = 4
## segment 36: lambda = 0.0002531, dev = 7.961e+04, npass = 4
## segment 37: lambda = 0.0002416, dev = 7.955e+04, npass = 4
## segment 38: lambda = 0.0002306, dev = 7.949e+04, npass = 4
## segment 39: lambda = 0.0002201, dev = 7.943e+04, npass = 4
## segment 40: lambda = 0.0002101, dev = 7.937e+04, npass = 4
## segment 41: lambda = 0.0002006, dev = 7.931e+04, npass = 4
## segment 42: lambda = 0.0001915, dev = 7.925e+04, npass = 4
## segment 43: lambda = 0.0001828, dev = 7.919e+04, npass = 4
## segment 44: lambda = 0.0001745, dev = 7.913e+04, npass = 4
## segment 45: lambda = 0.0001665, dev = 7.907e+04, npass = 4
## segment 46: lambda = 0.000159, dev = 7.901e+04, npass = 4
## segment 47: lambda = 0.0001517, dev = 7.895e+04, npass = 4
## segment 48: lambda = 0.0001448, dev = 7.889e+04, npass = 4
## segment 49: lambda = 0.0001382, dev = 7.883e+04, npass = 4
## segment 50: lambda = 0.000132, dev = 7.877e+04, npass = 4
## segment 51: lambda = 0.000126, dev = 7.871e+04, npass = 4
## segment 52: lambda = 0.0001202, dev = 7.865e+04, npass = 4
## segment 53: lambda = 0.0001148, dev = 7.86e+04, npass = 4
## segment 54: lambda = 0.0001096, dev = 7.854e+04, npass = 4
## segment 55: lambda = 0.0001046, dev = 7.848e+04, npass = 4
## segment 56: lambda = 9.983e-05, dev = 7.843e+04, npass = 4
## segment 57: lambda = 9.529e-05, dev = 7.838e+04, npass = 4
## segment 58: lambda = 9.096e-05, dev = 7.832e+04, npass = 4
## segment 59: lambda = 8.682e-05, dev = 7.827e+04, npass = 4
## segment 60: lambda = 8.288e-05, dev = 7.821e+04, npass = 4
## segment 61: lambda = 7.911e-05, dev = 7.816e+04, npass = 4
## segment 62: lambda = 7.552e-05, dev = 7.811e+04, npass = 4
## segment 63: lambda = 7.208e-05, dev = 7.806e+04, npass = 4
## segment 64: lambda = 6.881e-05, dev = 7.801e+04, npass = 4
## segment 65: lambda = 6.568e-05, dev = 7.795e+04, npass = 4
## segment 66: lambda = 6.269e-05, dev = 7.79e+04, npass = 4
## segment 67: lambda = 5.984e-05, dev = 7.786e+04, npass = 4
## segment 68: lambda = 5.712e-05, dev = 7.781e+04, npass = 4
## segment 69: lambda = 5.453e-05, dev = 7.776e+04, npass = 4
## segment 70: lambda = 5.205e-05, dev = 7.772e+04, npass = 3
## segment 71: lambda = 4.968e-05, dev = 7.767e+04, npass = 4
## segment 72: lambda = 4.743e-05, dev = 7.763e+04, npass = 3
## segment 73: lambda = 4.527e-05, dev = 7.758e+04, npass = 3
```

```
## segment 74: lambda = 4.321e-05, dev = 7.754e+04, npass = 3
## segment 75: lambda = 4.125e-05, dev = 7.75e+04, npass = 3
## segment 76: lambda = 3.937e-05, dev = 7.745e+04, npass = 3
## segment 77: lambda = 3.758e-05, dev = 7.741e+04, npass = 3
## segment 78: lambda = 3.588e-05, dev = 7.737e+04, npass = 3
## segment 79: lambda = 3.425e-05, dev = 7.733e+04, npass = 3
## segment 80: lambda = 3.269e-05, dev = 7.729e+04, npass = 3
## segment 81: lambda = 3.12e-05, dev = 7.725e+04, npass = 3
## segment 82: lambda = 2.978e-05, dev = 7.721e+04, npass = 3
## segment 83: lambda = 2.843e-05, dev = 7.717e+04, npass = 3
## segment 84: lambda = 2.714e-05, dev = 7.714e+04, npass = 3
## segment 85: lambda = 2.591e-05, dev = 7.71e+04, npass = 3
## segment 86: lambda = 2.473e-05, dev = 7.707e+04, npass = 3
## segment 87: lambda = 2.36e-05, dev = 7.703e+04, npass = 3
## segment 88: lambda = 2.253e-05, dev = 7.7e+04, npass = 3
## segment 89: lambda = 2.151e-05, dev = 7.697e+04, npass = 1
## segment 90: lambda = 2.053e-05, dev = 7.695e+04, npass = 1
## segment 91: lambda = 1.96e-05, dev = 7.691e+04, npass = 3
## segment 92: lambda = 1.871e-05, dev = 7.688e+04, npass = 3
## segment 93: lambda = 1.786e-05, dev = 7.685e+04, npass = 3
## segment 94: lambda = 1.704e-05, dev = 7.683e+04, npass = 1
## segment 95: lambda = 1.627e-05, dev = 7.68e+04, npass = 1
## segment 96: lambda = 1.553e-05, dev = 7.678e+04, npass = 1
## segment 97: lambda = 1.482e-05, dev = 7.676e+04, npass = 1
## segment 98: lambda = 1.415e-05, dev = 7.672e+04, npass = 3
## segment 99: lambda = 1.351e-05, dev = 7.67e+04, npass = 3
## segment 100: lambda = 1.289e-05, dev = 7.667e+04, npass = 3
```

```
plot(nhlreg)
```



Extracting Coefficents Let's grab the coefficients of each player

```
# AICc selection
Baicc <- coef(nhlreg)[colnames(player),]
```

First, a simple gut-check point: the intercept. This is the effect on odds that a goal is home rather than away, regardless of any info about what teams are playing or who is on ice. It's the home ice advantage!

We find that home-ice increases odds you've scored by 8%

```
100*(exp(coef(nhlreg)[1])-1)
```

```
## [1] 8.235434
```

Now, lets look at the player effects. The regression finds 646 significant player effects

```
sum(Baicc!=0)
```

```
## [1] 646
```

Here are the top 10 players

```
Baicc[order(Baicc, decreasing=TRUE)[1:10]]
```

```
## PETER_FORSBERG  TYLER_TOFFOLI   ONDREJ_PALAT ZIGMUND_PALFFY  SIDNEY_CROSBY
##      0.7548254      0.6292577      0.6284040      0.4426997      0.4131174
##   JOE_THORNTON  PAVEL_DATSYUK  LOGAN_COUTURE      ERIC_FEHR MARTIN_GELINAS
##      0.3837632      0.3761981      0.3682103      0.3677283      0.3577613
```

Here are the bottom 10

```
Baicc[order(Baicc)[1:10]]
```

```
##     TIM_TAYLOR  JOHN_MCCARTHY P. J._AXELSSON NICLAS_HAVELID    THOMAS_POCK
##     -0.8643214     -0.5651886     -0.4283811     -0.3854583     -0.3844128
##  MATHIEU_BIRON  CHRIS_DINGMAN   DARROLL_POWE RAITIS_IVANANS   RYAN_HOLLWEG
##     -0.3512101     -0.3342243     -0.3339906     -0.3129481     -0.2988769
```

Specifically, the model says, e.g., that whenever a goal is scored, Pittsburgh's odds of having scored (rather than scored on) increase by a 51% if Sidney Crosby is on the ice.

```
exp(Baicc["SIDNEY_CROSBY"])
```

```
## SIDNEY_CROSBY
##      1.511523
```

And the Blue Jackets (or Kings, pre 2011-12) odds of having scored drop by around 22% if Jack Johnson is on the ice.

```
exp(Baicc["JACK_JOHNSON"])
```

```
## JACK_JOHNSON
##    0.7813488
```

hockey fans among you may feel free to comment in much more detail.

Standardizing vs Not Standardizing Data

Without `standardize=FALSE`, you'd be multiplying the penalty for each coefficient (player effect) by that player's standard deviation in onice. The players with big SD in onice are guys who play a lot. Players with small SD are those who play little (almost all zeros). So weighting penalty by SD in this case is exactly what you don't want: a bigger penalty for people with many minutes on ice, a smaller penalty for those who seldom play. Indeed, running the regression without `standardize=FALSE` leads to a bunch of farm teamers coming up tops.

```
nhlreg.std <-  gamlr(x, y,
    free=1:(ncol(config)+ncol(team)), family="binomial")
Bstd <- coef(nhlreg.std)[colnames(player),]
Bstd[order(Bstd, decreasing=TRUE)[1:10]]
```

```
##        JEFF_TOMS       RYAN_KRAFT     COLE_JARRETT   TOMAS_POPPERLE
##        1.7380706        1.4826419        1.2119318        1.1107806
##   DAVID_LIFFITON ALEXEY_MARCHENKO     ERIC_SELLECK      MIKE_MURPHY
##        1.0974872        1.0297324        1.0060015        0.9600939
##       DAVID_GOVE       TOMAS_KANA
##        0.9264895        0.8792802
```

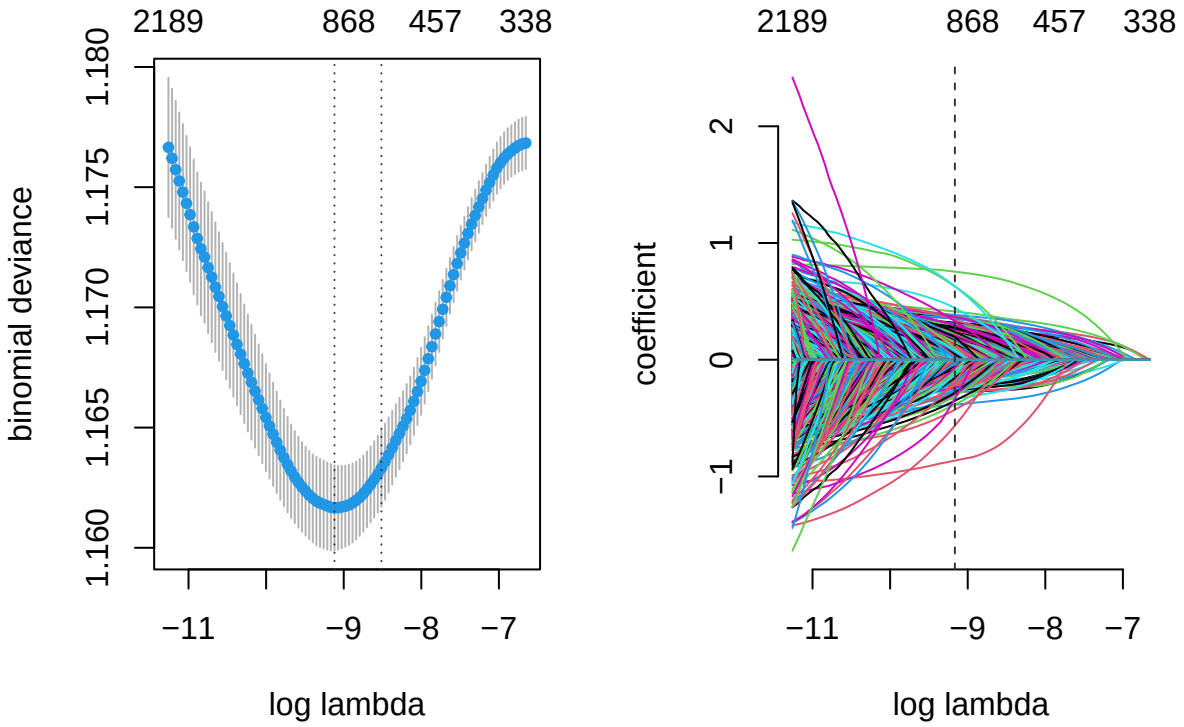
NOTE: this is an exceptional case! You almost always want standardize=TRUE.

Cross-Validation

```
cv.nhlreg <- cv.gamlr(x, y,
    free=1:(ncol(config)+ncol(team)),
    family="binomial", verb=TRUE, standardize=FALSE)
```

```
## fold 1,2,3,4,5,done.
```

```
par(mfrow=c(1,2))
plot(cv.nhlreg)
plot(cv.nhlreg$gamlr) ## cv.gamlr has included a gamlr object into cv.nhlreg
```



Plotting Results

Lambda Selection

```
log(nhlreg$lambda[which.min(AICc(nhlreg))])
```

```
##     seg55
## -9.165555
```

```
log(nhlreg$lambda[which.min(AIC(nhlreg))])
```

```
##     seg55
## -9.165555
```

```
log(nhlreg$lambda[which.min(BIC(nhlreg))])
```

```
##      seg1
## -6.653644
```

```
log(cv.nhlreg$lambda.min)
```

```
## [1] -9.119038
```

```
log(cv.nhlreg$lambda.1se)
```

```
## [1] -8.514319
```

AIC and AICc give exactly the same answer here (n»df) and both are close to the `cv.min` answer.

```
Bcvmin <- coef(cv.nhlreg, select="min")[colnames(player),]
sum(Bcvmin!=0) # around 600
```

```
## [1] 622
```

```
exp(sort(Bcvmin,decreasing=TRUE)[1:10]) # similar top 10
```

```
## PETER_FORSBERG   ONDREJ_PALAT  TYLER_TOFFOLI ZIGMUND_PALFFY  SIDNEY_CROSBY
##       2.119080       1.834845       1.830916       1.531845       1.506432
##   JOE_THORNTON  PAVEL_DATSYUK  LOGAN_COUTURE      ERIC_FEHR MARTIN_GELINAS
##       1.464191       1.449265       1.437288       1.435486       1.421528
```

Both AIC and AICc are trying to approximate the OOS deviance (MSE here). Thus the lambdas at minimum AIC and AICc values are estimates of the lambda which minimizes OOS error – the same thing targeted with the cv.min rule. Also, in this case, the degrees of freedom are low enough relative to 'n' that AIC works fine, and gives an answer close to AICc.

The 1se rule accounts for uncertainty about OOS error, and thus chooses a simpler model.

```
Bcv1se <- coef(cv.nhlreg)[colnames(player),]
```

Even though log lambdas are close, df drops by 1/2

```
sum(Bcv1se!=0) # only around 300
```

```
## [1] 341
```

Top 10 changes a bit as well

```
exp(sort(Bcv1se,decreasing=TRUE)[1:10]) # top 10 changes a bit
```

```
##    PETER_FORSBERG     SIDNEY_CROSBY      JOE_THORNTON     PAVEL_DATSYUK
##          1.973963          1.422968          1.403477          1.354780
##   ALEXANDER_SEMIN     LADISLAV_NAGY LUBOMIR_VISNOVSKY      MATT_MOULSON
##          1.351025          1.334468          1.334158          1.323587
##     LOGAN_COUTURE     ALEX_OVECHKIN
##          1.315394          1.311600
```

BIC is way more (overly I think) conservative than all these options.

```
Bbic <- coef(nhlreg,select=which.min(BIC(nhlreg)))[colnames(player),]
sum(Bbic!=0) # zero!  Nobody is different from average according to BIC
```

```
## [1] 0
```

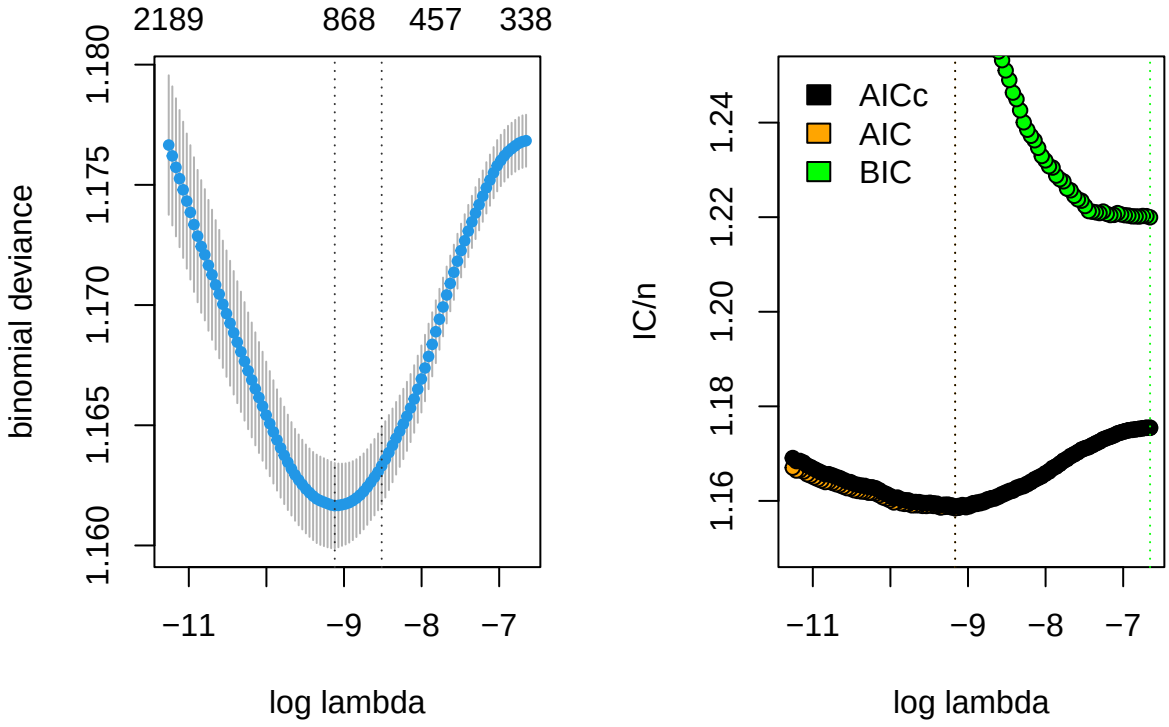
The BIC is trying to find lambda with highest probability of having the minimum OOS error, which is subtly different than finding the lambda corresponding to lowest expected OOS error. For example, if there is more uncertainty about OOS error at the lambda with min expectation, then it could be that another value with higher expected error but lower uncertainty around this value will have a higher probability of being best. In this case, the BIC says there is much uncertainty at everything other than the null model, so that the null model ends up highest probability of being best.

As an aside: note that the null model here is not just an intercept, but rather includes onice configuration info along with information about the team and season. So the BIC is not saying that no players matter, but rather that it cannot confidently tell them apart from their team's average level of play in a given season.

Comparing Model Selections

Finally, some plots to compare model selections

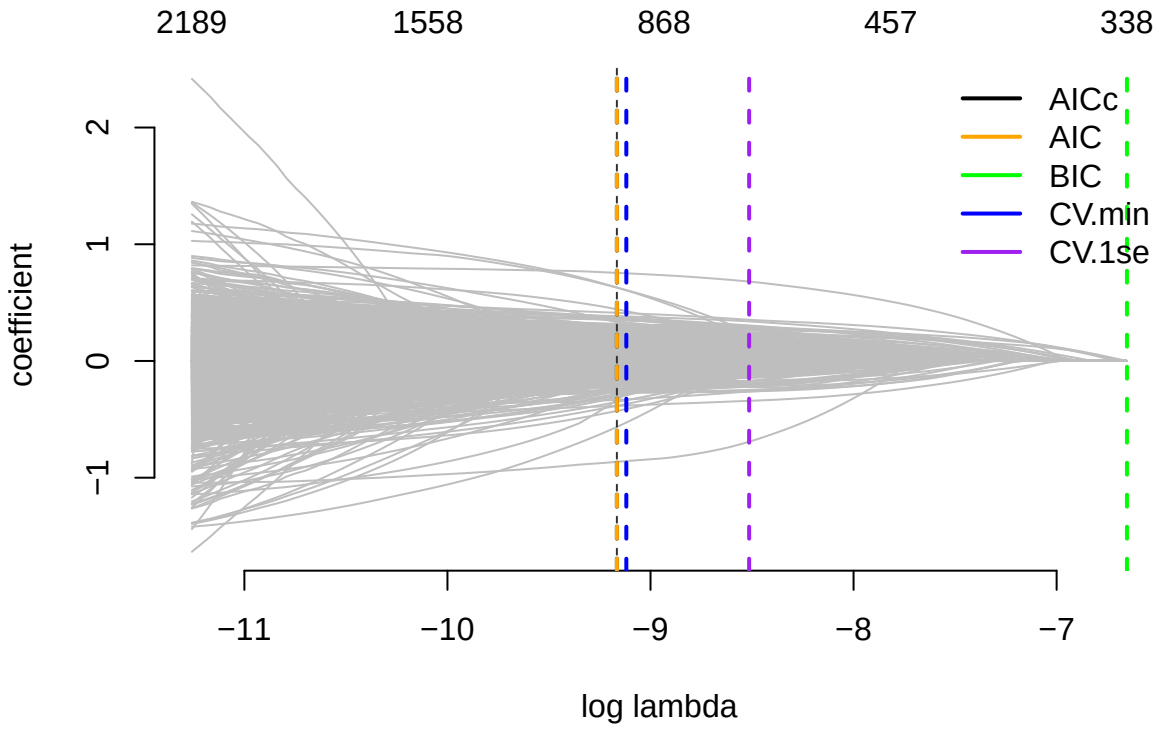
```
ll <- log(nhlreg$lambda) ## the sequence of lambdas
n <- nrow(goal)
par(mfrow=c(1,2))
plot(cv.nhlreg)
plot(ll, AIC(nhlreg)/n,
    xlab="log lambda", ylab="IC/n", pch=21, bg="orange",ylim = c(1.15,1.25))
abline(v=ll[which.min(AIC(nhlreg))], col="orange", lty=3)
points(ll, AICc(nhlreg)/n, pch=21, bg="black")
abline(v=ll[which.min(AICc(nhlreg))], col="black", lty=3)
points(ll, BIC(nhlreg)/n, pch=21, bg="green")
abline(v=ll[which.min(BIC(nhlreg))], col="green", lty=3)

legend("topleft", bty="n",
    fill=c("black","orange","green"),legend=c("AICc","AIC","BIC"))
```



Plot all the answers along the path

```
par(mfrow=c(1,1))
plot(nhlreg, col="grey")
abline(v=ll[which.min(AICc(nhlreg))], col="black", lty=2, lwd=2)
abline(v=ll[which.min(AIC(nhlreg))], col="orange", lty=2, lwd=2)
abline(v=ll[which.min(BIC(nhlreg))], col="green", lty=2, lwd=2)
abline(v=log(cv.nhlreg$lambda.min), col="blue", lty=2, lwd=2)
abline(v=log(cv.nhlreg$lambda.1se), col="purple", lty=2, lwd=2)
legend("topright", bty="n", lwd=2,
    col=c("black","orange","green","blue","purple"),
```

```
    legend=c("AICc","AIC","BIC","CV.min","CV.1se"))
```



Browsing History

The table has three columns: [machine] id, site [id], [# of] visits

Data Mangement

```
web <- read.csv("browser-domains.csv")
```

Read in the actual website names and relabel site factor

```
sitenames <- scan("browser-sites.txt", what="character")
web$site <- factor(web$site, levels=1:length(sitenames), labels=sitenames)
```

also factor machine id

```
web$id <- factor(web$id, levels=1:length(unique(web$id)))
```

Let's get total visits per-machine and % of time on each site `tapply(a,b,c)` does `c(a)` for every level of factor `b`.

```
machinetotals <- as.vector(tapply(web$visits,web$id,sum))
visitpercent <- 100*web$visits/machinetotals[web$id]
```

Since there are many sites that some people do not visit, our dataframe could blow out of proportion. We will first need to generate a sparse matrix.

```
xweb <- sparseMatrix(
    i=as.numeric(web$id), j=as.numeric(web$site), x=visitpercent,
    dims=c(nlevels(web$id),nlevels(web$site)),
    dimnames=list(id=levels(web$id), site=levels(web$site)))
```

We can extract values from the sparse matrix. Ex: What sites did household 1 visit?

```
xweb[1, xweb[1,]!=0]
```

```
##                  atdmt.com                  yahoo.com
##                 4.05202601                11.85592796
##                    msn.com                 google.com
##                 0.25012506                 6.52826413
##                    aol.com          questionmarket.com
##                 0.15007504                 1.35067534
##   googlesyndication.com-o02             casalemedia.com
##                 3.40170085                 0.82541271
##             mywebsearch.com                myspace.com
##                 0.10005003                 0.05002501
##               pointroll.com            yieldmanager.com
##                 0.72536268                 0.65032516
##                   live.com              mediaplex.com
##                 0.02501251                 0.32516258
##          precisionclick.com            tribalfusion.com
##                 0.15007504                 0.52526263
##        insightexpressai.com               trafficmp.com
##                 0.65032516                 0.30015008
##                   ebay.com              realmedia.com
##                 1.97598799                 0.52526263
##                   zedo.com             advertising.com
##                 0.62531266                 0.40020010
##               microsoft.com                 hotbar.com
##                 0.57528764                 0.02501251
##              adrevolver.com                    ru4.com
##                 0.35017509                 0.27513757
##                 nextag.com               overture.com
##                 1.67583792                 0.07503752
##                starware.com       relevantknowledge.com
##                 0.10005003                 2.37618809
##                  myway.com          partner2profit.com
##                 1.12556278                 2.92646323
##                  ditto.com                kanoodle.com
##                 0.07503752                 0.20010005
##             ebayobjects.com                comcast.net
##                 1.35067534                 0.07503752
##               fastclick.net                adbrite.com
##                 0.57528764                 0.50025013
##         vpptechnologies.com           specificclick.net
##                 0.05002501                 0.32516258
##             serving-sys.com                weather.com
##                 0.40020010                 0.15007504
##                adserver.com                     go.com
##                 0.15007504                 0.42521261
##              intellisrv.net                   dell.com
```

```
##                 0.20010005                4.77738869
##                    cnn.com               burstnet.com
##                 0.05002501                0.45022511
##             adknowledge.com          funwebproducts.com
##                 0.27513757                0.15007504
##                  belnk.com               netscape.com
##                 0.02501251                0.05002501
##                   real.com             liveperson.net
##                 0.02501251                0.05002501
##                adsonar.com              euroclick.com
##                 0.15007504                0.07503752
##                    m7z.net              mywebface.com
##                 0.20010005                0.02501251
##                 tacoda.net                unicast.com
##                 0.85042521                0.10005003
##             personalweb.com              earthlink.net
##                 0.20010005                0.07503752
##                 falkag.net                amazon.com
##                 0.07503752                1.05052526
##                dellfix.com                ysbweb.com
##                 3.90195098                0.02501251
##         googleadservices.com                  qnsr.com
##                 0.25012506                2.17608804
##                revenue.net              addynamix.com
##                 0.07503752                0.12506253
##                    cox.net                akamai.net
##                13.38169085                0.02501251
##            admarketplace.net             amazon.com-o01
##                 0.37518759                0.40020010
##            opinionsquare.com             interclick.com
##                 0.05002501                0.07503752
##                 go.com-o04                   ask.com
##                 0.02501251                0.60030015
##              contextweb.com             intellitxt.com
##                 0.25012506                0.22511256
##                  yceml.net                 about.com
##                 0.12506253                0.40020010
##                youtube.com              wikipedia.org
##                 0.07503752                1.32566283
##             windowsmedia.com              foxsports.com
##                 0.07503752                0.02501251
##                walmart.com             partypoker.com
##                 0.07503752                0.05002501
##              mymailstamp.com                windows.com
##                 0.02501251                0.02501251
##               musicmatch.com           specificmedia.com
##                 1.60080040                0.05002501
##                 paypal.com             monster.com-o01
##                 0.52526263                0.17508754
##              information.com                   2o7.net
##                 0.12506253                0.42521261
##                ebayrtm.com                 match.com
##                 0.15007504                0.02501251
##             screensavers.com           smileycentral.com
```

```
##                0.05002501                1.30065033
##          careerbuilder.com                   mlb.com
##                0.10005003                0.05002501
##           searchignite.com           thinktarget.com
##                0.30015008                0.12506253
##                dotomi.com               blogspot.com
##                0.02501251                0.25012506
##                 chase.com                vonage.com
##                0.02501251                0.02501251
##             compuserve.com                 lycos.com
##                0.07503752                0.15007504
##                 azjmp.com              tmcs.net-o01
##                0.07503752                0.05002501
##           exitexchange.com                  imdb.com
##                0.05002501                0.07503752
##                 adecn.com                go.com-o03
##                0.10005003                0.07503752
##                excite.com                imiclk.com
##                0.02501251                0.02501251
##              bloglines.com               foxnews.com
##                0.55027514                0.02501251
##            marketwatch.com               monster.com
##                0.25012506                0.02501251
##                 adobe.com               chitika.net
##                0.05002501                0.02501251
##              geocities.com       sharewareonline.com
##                0.10005003                0.07503752
##                 llnwd.net                 oingo.com
##                0.17508754                0.07503752
##          lightningcast.net                   xolox.nl
##                0.02501251                0.12506253
##                aavalue.com             bluestreak.com
##                0.02501251                0.02501251
##                  usps.com                 bbc.co.uk
##                0.10005003                0.05002501
##                   att.net            alumnigroup.org
##                0.02501251                0.10005003
##        verizonwireless.com               webshots.com
##                0.17508754                0.12506253
##               bestbuy.com               paypopup.com
##                0.05002501                0.02501251
##            cheaptickets.com                nytimes.com
##                0.02501251                0.02501251
##               cpmstar.com           ticketmaster.com
##                0.02501251                0.05002501
##                tripod.com                   buy.com
##                0.10005003                0.05002501
##             contextuads.com             hbmediapro.com
##                0.02501251                0.02501251
##               usatoday.com               adoutput.com
##                0.05002501                0.02501251
##                   sbc.com           lowermybills.com
##                0.02501251                0.02501251
##               clicksor.com                 legacy.com
```

```
##                  0.02501251                  0.02501251
##                    hgtv.com                sprintpcs.com
##                  0.25012506                  0.02501251
##                   webmd.com              clearchannel.com
##                  0.02501251                  0.05002501
##                    tmcs.net                       tx.us
##                  0.05002501                  0.05002501
##                    sony.com                citysearch.com
##                  0.02501251                  0.02501251
##                connextra.com                adlegend.com
##                  0.02501251                  0.02501251
##                   sears.com                       fl.us
##                  0.02501251                  0.02501251
##              coolsavings.com     liquidmedianetworks.com
##                  0.02501251                  0.02501251
##             perfectmatch.com               stockgroup.com
##                  0.25012506                  0.02501251
## consumerpromotioncenter.com                  kontera.com
##                  0.02501251                  0.02501251
##              circuitcity.com                    yimg.com
##                  0.02501251                  0.05002501
##             wunderground.com                       ca.us
##                  0.02501251                  0.02501251
##           washingtonpost.com                adjuggler.com
##                  0.07503752                  0.12506253
##              scripps.com-o01                     irs.gov
##                  0.12506253                  0.05002501
##                     nih.gov                     ign.com
##                  0.15007504                  0.02501251
##               proficient.com                  checkm8.com
##                  0.02501251                  0.02501251
##                   oasei.com                  passion.com
##                  0.02501251                  0.02501251
##      trafficmarketplace.com                   flickr.com
##                  0.02501251                  0.07503752
##                  andale.com              entrepreneur.com
##                  0.20010005                  0.02501251
##                webmd.com-o01                bbeplayer.com
##                  0.02501251                  0.02501251
##                 204.95.60.12                  napster.com
##                  0.02501251                  0.42521261
##                   msnbc.com              linkexchange.com
##                  0.02501251                  0.02501251
##          searchmarketing.com                 angelfire.com
##                  0.05002501                  0.05002501
##                  spynet.com                macromedia.com
##                  1.50075038                  0.05002501
##                      ed.gov               frontiernet.net
##                  0.17508754                  0.02501251
##                       nc.us          ticketmaster.com-o01
##                  0.05002501                  0.02501251
##                  flowgo.com                  oddcast.com
##                  0.12506253                  0.05002501
##                 answers.com                   lowes.com
```

```
##                0.10005003                0.05002501
##                  shop.com             orchardbank.com
##                0.02501251                0.02501251
##             google.com-o03                   cbs.com
##                0.12506253                0.05002501
##              bannerspace.com             technorati.com
##                0.05002501                0.05002501
##             ebaumsworld.com               gamespot.com
##                0.05002501                0.07503752
##                  bebo.com                subsag.com
##                0.07503752                0.05002501
##                 local.com                  gms1.net
##                0.02501251                0.02501251
##           gannettonline.com           superbrewards.com
##                0.07503752                0.02501251
##               ugo.com-o02               comcast.com
##                0.02501251                0.02501251
##                   sun.com           gamespot.com-o01
##                0.05002501                0.05002501
##              youravon.com                 yahoo.net
##                0.05002501                0.02501251
##                go.com-o02      deepnetexplorer.co.uk
##                0.05002501                0.10005003
##                 break.com               hallmark.com
##                0.20010005                0.02501251
##             pro-market.net                 mate1.com
##                0.05002501                0.02501251
##                tickle.com            bluemountain.com
##                0.05002501                0.07503752
##               blogger.com               xmradio.com
##                0.05002501                0.02501251
##              ancestry.com             imixserver.com
##                0.02501251                0.10005003
##            pricegrabber.com               limewire.com
##                0.02501251                0.05002501
##          freelaptop4you.com                typepad.com
##                0.05002501                0.02501251
##             allrecipes.com               eversave.com
##                0.02501251                0.02501251
##                boston.com      ibanking-services.com
##                0.02501251                0.02501251
##              custhelp.com                     mn.us
##                0.02501251                0.02501251
##              gamefaqs.com   premiumproductsonline.com
##                0.02501251                0.02501251
##                    tv.com               65.115.67.11
##                0.02501251                0.02501251
##              snapfish.com           artistdirect.com
##                0.02501251                0.02501251
##             scottrade.com                 rootv.com
##                0.02501251                0.02501251
##               webstat.com           inklineglobal.com
##                0.02501251                0.02501251
##                  seeq.com           sonypictures.com
```

```
##                  0.07503752                  0.02501251
##                   slate.com              123greetings.com
##                  0.02501251                  0.10005003
##                  blinko.com                    turn.com
##                  0.02501251                  0.02501251
##           householdbank.com                  active.com
##                  0.02501251                  0.02501251
##                michigan.gov              dada-mobile.net
##                  0.02501251                  0.02501251
##                pennyweb.com                     tmz.com
##                  0.02501251                  0.02501251
##              blogrolling.com                samsclub.com
##                  0.15007504                  0.02501251
##                 toseeka.com                 gamespy.com
##                  0.02501251                  0.02501251
##                cafepress.com                      ea.com
##                  0.22511256                  0.07503752
##                petfinder.com       websourcedtraffic.com
##                  0.02501251                  0.05002501
##         trafficexplorer.com           freegiftworld.com
##                  0.02501251                  0.02501251
##                  zap2it.com              gadgetcity.com
##                  0.02501251                  0.02501251
##                    army.mil           yourgiftcards.com
##                  0.02501251                  0.02501251
##              statcounter.com                     ajc.com
##                  0.02501251                  0.02501251
##            jacquielawson.com                  atomz.com
##                  0.05002501                  0.05002501
##                   topix.net                mybloglog.com
##                  0.07503752                  0.02501251
##                    away.com                    grab.com
##                  0.02501251                  0.02501251
##                     pbs.org                business.com
##                  0.12506253                  0.02501251
##           myinsiderdeals.com                homegain.com
##                  0.02501251                  0.02501251
##             findarticles.com              cursormania.com
##                  0.10005003                  0.02501251
##     thebreastcancersite.com                      cox.com
##                  0.05002501                  0.02501251
```

Read in the spending data

```
yspend <- read.csv("browser-totalspend.csv", row.names=1)  # use 1st column as row names
yspend <- as.matrix(yspend) ## good practice to move from dataframe to matrix
```

LASSO

```
spender <- gamlr(xweb, log(yspend), verb=TRUE)
```

```
## *** n=10000 observations and p=1000 covariates ***
## segment 1: lambda = 0.2325, dev = 2.783e+04, npass = 0
## segment 2: lambda = 0.2219, dev = 2.778e+04, npass = 3
## segment 3: lambda = 0.2118, dev = 2.774e+04, npass = 3
```

```
## segment 4: lambda = 0.2022, dev = 2.77e+04, npass = 3
## segment 5: lambda = 0.193, dev = 2.766e+04, npass = 3
## segment 6: lambda = 0.1842, dev = 2.76e+04, npass = 3
## segment 7: lambda = 0.1759, dev = 2.753e+04, npass = 4
## segment 8: lambda = 0.1679, dev = 2.744e+04, npass = 5
## segment 9: lambda = 0.1602, dev = 2.733e+04, npass = 5
## segment 10: lambda = 0.153, dev = 2.72e+04, npass = 5
## segment 11: lambda = 0.146, dev = 2.708e+04, npass = 5
## segment 12: lambda = 0.1394, dev = 2.697e+04, npass = 5
## segment 13: lambda = 0.133, dev = 2.687e+04, npass = 5
## segment 14: lambda = 0.127, dev = 2.676e+04, npass = 5
## segment 15: lambda = 0.1212, dev = 2.663e+04, npass = 5
## segment 16: lambda = 0.1157, dev = 2.65e+04, npass = 5
## segment 17: lambda = 0.1104, dev = 2.639e+04, npass = 5
## segment 18: lambda = 0.1054, dev = 2.628e+04, npass = 5
## segment 19: lambda = 0.1006, dev = 2.616e+04, npass = 5
## segment 20: lambda = 0.09606, dev = 2.605e+04, npass = 5
## segment 21: lambda = 0.09169, dev = 2.595e+04, npass = 5
## segment 22: lambda = 0.08753, dev = 2.584e+04, npass = 5
## segment 23: lambda = 0.08355, dev = 2.574e+04, npass = 4
## segment 24: lambda = 0.07975, dev = 2.564e+04, npass = 4
## segment 25: lambda = 0.07613, dev = 2.553e+04, npass = 4
## segment 26: lambda = 0.07267, dev = 2.543e+04, npass = 4
## segment 27: lambda = 0.06936, dev = 2.532e+04, npass = 4
## segment 28: lambda = 0.06621, dev = 2.522e+04, npass = 4
## segment 29: lambda = 0.0632, dev = 2.511e+04, npass = 4
## segment 30: lambda = 0.06033, dev = 2.501e+04, npass = 4
## segment 31: lambda = 0.05759, dev = 2.491e+04, npass = 4
## segment 32: lambda = 0.05497, dev = 2.481e+04, npass = 4
## segment 33: lambda = 0.05247, dev = 2.47e+04, npass = 4
## segment 34: lambda = 0.05009, dev = 2.46e+04, npass = 4
## segment 35: lambda = 0.04781, dev = 2.449e+04, npass = 5
## segment 36: lambda = 0.04564, dev = 2.438e+04, npass = 5
## segment 37: lambda = 0.04356, dev = 2.427e+04, npass = 6
## segment 38: lambda = 0.04158, dev = 2.416e+04, npass = 6
## segment 39: lambda = 0.03969, dev = 2.405e+04, npass = 6
## segment 40: lambda = 0.03789, dev = 2.395e+04, npass = 6
## segment 41: lambda = 0.03617, dev = 2.384e+04, npass = 6
## segment 42: lambda = 0.03452, dev = 2.374e+04, npass = 6
## segment 43: lambda = 0.03295, dev = 2.365e+04, npass = 6
## segment 44: lambda = 0.03146, dev = 2.355e+04, npass = 5
## segment 45: lambda = 0.03003, dev = 2.346e+04, npass = 5
## segment 46: lambda = 0.02866, dev = 2.337e+04, npass = 5
## segment 47: lambda = 0.02736, dev = 2.328e+04, npass = 5
## segment 48: lambda = 0.02611, dev = 2.319e+04, npass = 5
## segment 49: lambda = 0.02493, dev = 2.311e+04, npass = 5
## segment 50: lambda = 0.02379, dev = 2.302e+04, npass = 5
## segment 51: lambda = 0.02271, dev = 2.293e+04, npass = 5
## segment 52: lambda = 0.02168, dev = 2.285e+04, npass = 5
## segment 53: lambda = 0.0207, dev = 2.277e+04, npass = 5
## segment 54: lambda = 0.01975, dev = 2.269e+04, npass = 5
## segment 55: lambda = 0.01886, dev = 2.261e+04, npass = 4
## segment 56: lambda = 0.018, dev = 2.253e+04, npass = 4
## segment 57: lambda = 0.01718, dev = 2.245e+04, npass = 4
```

```
## segment 58: lambda = 0.0164, dev = 2.238e+04, npass = 4
## segment 59: lambda = 0.01566, dev = 2.23e+04, npass = 4
## segment 60: lambda = 0.01494, dev = 2.223e+04, npass = 4
## segment 61: lambda = 0.01426, dev = 2.216e+04, npass = 5
## segment 62: lambda = 0.01362, dev = 2.209e+04, npass = 4
## segment 63: lambda = 0.013, dev = 2.203e+04, npass = 4
## segment 64: lambda = 0.01241, dev = 2.197e+04, npass = 4
## segment 65: lambda = 0.01184, dev = 2.191e+04, npass = 4
## segment 66: lambda = 0.0113, dev = 2.186e+04, npass = 4
## segment 67: lambda = 0.01079, dev = 2.18e+04, npass = 4
## segment 68: lambda = 0.0103, dev = 2.175e+04, npass = 6
## segment 69: lambda = 0.009832, dev = 2.17e+04, npass = 4
## segment 70: lambda = 0.009385, dev = 2.165e+04, npass = 6
## segment 71: lambda = 0.008959, dev = 2.16e+04, npass = 7
## segment 72: lambda = 0.008551, dev = 2.155e+04, npass = 7
## segment 73: lambda = 0.008163, dev = 2.151e+04, npass = 6
## segment 74: lambda = 0.007792, dev = 2.147e+04, npass = 6
## segment 75: lambda = 0.007438, dev = 2.143e+04, npass = 6
## segment 76: lambda = 0.007099, dev = 2.139e+04, npass = 6
## segment 77: lambda = 0.006777, dev = 2.136e+04, npass = 6
## segment 78: lambda = 0.006469, dev = 2.132e+04, npass = 6
## segment 79: lambda = 0.006175, dev = 2.129e+04, npass = 7
## segment 80: lambda = 0.005894, dev = 2.126e+04, npass = 7
## segment 81: lambda = 0.005626, dev = 2.122e+04, npass = 10
## segment 82: lambda = 0.00537, dev = 2.119e+04, npass = 10
## segment 83: lambda = 0.005126, dev = 2.117e+04, npass = 9
## segment 84: lambda = 0.004893, dev = 2.114e+04, npass = 9
## segment 85: lambda = 0.004671, dev = 2.112e+04, npass = 9
## segment 86: lambda = 0.004459, dev = 2.109e+04, npass = 8
## segment 87: lambda = 0.004256, dev = 2.107e+04, npass = 8
## segment 88: lambda = 0.004063, dev = 2.105e+04, npass = 8
## segment 89: lambda = 0.003878, dev = 2.103e+04, npass = 7
## segment 90: lambda = 0.003702, dev = 2.102e+04, npass = 7
## segment 91: lambda = 0.003533, dev = 2.1e+04, npass = 7
## segment 92: lambda = 0.003373, dev = 2.098e+04, npass = 7
## segment 93: lambda = 0.00322, dev = 2.097e+04, npass = 6
## segment 94: lambda = 0.003073, dev = 2.095e+04, npass = 6
## segment 95: lambda = 0.002934, dev = 2.094e+04, npass = 6
## segment 96: lambda = 0.0028, dev = 2.093e+04, npass = 6
## segment 97: lambda = 0.002673, dev = 2.091e+04, npass = 6
## segment 98: lambda = 0.002551, dev = 2.09e+04, npass = 5
## segment 99: lambda = 0.002435, dev = 2.089e+04, npass = 5
## segment 100: lambda = 0.002325, dev = 2.088e+04, npass = 5
```

```
summary(spender)
```

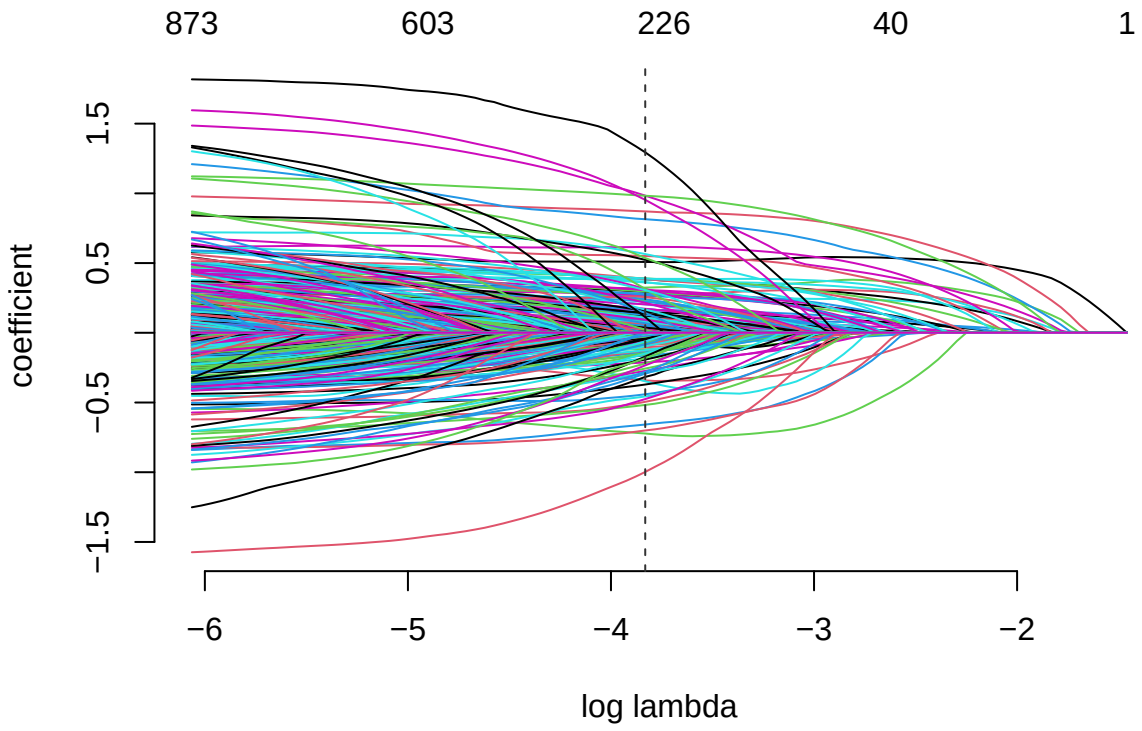
```
##
## gaussian gamlr with 1000 inputs and 100 segments.
```

```
##             lambda par  df          r2      aicc
## seg1   0.232476092   1   1 0.000000000 10236.678
## seg2   0.221909695   2   2 0.001725303 10221.411
## seg3   0.211823557   2   2 0.003297334 10205.651
## seg4   0.202195849   2   2 0.004729710 10191.270
## seg5   0.193005736   2   2 0.006034838 10178.148
```

```
## seg6   0.184233327   3   3 0.008059419 10159.760
## seg7   0.175859638   5   5 0.010759747 10136.503
## seg8   0.167866546   7   7 0.013978241 10107.920
## seg9   0.160236753  10  10 0.018051482 10072.536
## seg10  0.152953745  11  11 0.022568088 10028.438
## seg11  0.146001761  12  12 0.026834473  9986.698
## seg12  0.139365756  12  12 0.030824198  9945.617
## seg13  0.133031367  13  13 0.034577333  9908.822
## seg14  0.126984886  16  16 0.038557614  9873.526
## seg15  0.121213227  18  18 0.043190423  9829.238
## seg16  0.115703899  18  18 0.047623601  9782.797
## seg17  0.110444978  18  18 0.051662879  9740.294
## seg18  0.105425083  21  21 0.055664541  9704.032
## seg19  0.100633351  22  22 0.059819719  9661.943
## seg20  0.096059410  24  24 0.063739653  9624.181
## seg21  0.091693361  25  25 0.067595723  9584.920
## seg22  0.087525756  28  28 0.071296509  9551.183
## seg23  0.083547576  31  31 0.074986542  9517.407
## seg24  0.079750209  31  31 0.078668776  9477.520
## seg25  0.076125439  37  37 0.082487125  9448.073
## seg26  0.072665421  40  40 0.086306440  9412.407
## seg27  0.069362665  42  42 0.090037319  9375.524
## seg28  0.066210025  49  49 0.093679230  9349.550
## seg29  0.063200677  54  54 0.097580298  9316.519
## seg30  0.060328109  54  54 0.101267227  9275.580
## seg31  0.057586103  58  58 0.104923128  9242.910
## seg32  0.054968726  63  63 0.108525962  9212.700
## seg33  0.052470312  68  68 0.112302698  9180.378
## seg34  0.050085456  79  79 0.116044277  9160.469
## seg35  0.047808994  87  87 0.119877269  9133.283
## seg36  0.045636002  92  92 0.123789610  9098.915
## seg37  0.043561775 101 101 0.127898571  9070.264
## seg38  0.041581825 108 108 0.131909537  9038.465
## seg39  0.039691867 115 115 0.135762390  9008.302
## seg40  0.037887811 122 122 0.139472471  8979.620
## seg41  0.036165751 130 130 0.143148598  8953.221
## seg42  0.034521962 143 143 0.146752614  8937.799
## seg43  0.032952885 150 150 0.150323859  8910.277
## seg44  0.031455126 155 155 0.153730158  8880.420
## seg45  0.030025442 165 165 0.157075185  8861.473
## seg46  0.028660739 176 176 0.160363792  8845.155
## seg47  0.027358064 186 186 0.163519645  8828.244
## seg48  0.026114598 196 196 0.166631375  8811.763
## seg49  0.024927649 208 208 0.169708540  8799.773
## seg50  0.023794649 226 226 0.172763466  8800.530
## seg51  0.022713146 238 238 0.175884937  8787.882
## seg52  0.021680798 249 249 0.178885990  8774.514
## seg53  0.020695372 270 270 0.181813234  8783.073
## seg54  0.019754736 286 286 0.184724207  8781.291
## seg55  0.018856853 303 303 0.187588758  8782.192
## seg56  0.017999780 319 319 0.190425860  8781.300
## seg57  0.017181662 336 336 0.193222530  8783.040
## seg58  0.016400729 353 353 0.195936722  8785.814
## seg59  0.015655290 367 367 0.198593764  8782.848
```

```
## seg60  0.014943733 387 387 0.201160348  8793.971
## seg61  0.014264517 402 402 0.203659807  8795.152
## seg62  0.013616173 412 412 0.206047391  8786.860
## seg63  0.012997297 431 431 0.208312775  8799.709
## seg64  0.012406550 444 444 0.210482821  8800.697
## seg65  0.011842653 463 463 0.212587024  8815.711
## seg66  0.011304386 478 478 0.214637570  8822.674
## seg67  0.010790584 492 492 0.216619355  8828.339
## seg68  0.010300136 505 505 0.218518166  8832.874
## seg69  0.009831979 513 513 0.220339652  8827.303
## seg70  0.009385100 531 531 0.222145966  8844.187
## seg71  0.008958533 545 545 0.223896744  8852.932
## seg72  0.008551354 559 559 0.225538511  8863.126
## seg73  0.008162681 571 571 0.227089202  8870.047
## seg74  0.007791675 586 586 0.228546255  8884.979
## seg75  0.007437531 603 603 0.229936408  8905.381
## seg76  0.007099484 618 618 0.231283973  8921.898
## seg77  0.006776801 638 638 0.232580144  8950.568
## seg78  0.006468785 646 646 0.233810612  8952.794
## seg79  0.006174769 662 662 0.235015469  8973.696
## seg80  0.005894116 675 675 0.236187355  8988.227
## seg81  0.005626220 689 689 0.237333116  9005.467
## seg82  0.005370499 703 703 0.238393719  9023.901
## seg83  0.005126402 714 714 0.239395276  9036.228
## seg84  0.004893399 721 721 0.240338414  9040.070
## seg85  0.004670986 730 730 0.241208491  9049.538
## seg86  0.004458683 738 738 0.242015972  9057.529
## seg87  0.004256029 747 747 0.242766854  9068.623
## seg88  0.004062586 760 760 0.243477983  9089.641
## seg89  0.003877935 773 773 0.244149588  9111.259
## seg90  0.003701677 786 786 0.244792248  9133.338
## seg91  0.003533430 800 800 0.245411685  9158.168
## seg92  0.003372830 809 809 0.245997670  9171.689
## seg93  0.003219530 820 820 0.246538742  9190.588
## seg94  0.003073197 827 827 0.247066267  9200.212
## seg95  0.002933516 830 830 0.247559945  9200.787
## seg96  0.002800183 841 841 0.248026092  9220.787
## seg97  0.002672910 849 849 0.248465108  9234.039
## seg98  0.002551422 855 855 0.248856592  9243.170
## seg99  0.002435456 867 867 0.249232707  9266.900
## seg100 0.002324761 873 873 0.249589392  9276.545
```

```
plot(spender) ## path plot
```



Extracting and observing coefficients

DO NOT GO TO THESE LINKS

```
B <- coef(spender) ## the coefficients selected under AICc

B <- B[-1,] # drop intercept and remove STM formatting
B[which.min(B)] ## low spenders spend a lot of time here
```

```
## cursormania.com
##       -0.998143
```

```
B[which.max(B)] ## big spenders hang out here
```

```
## shopyourbargain.com
##            1.294246
```

```
coef(spender, select=which.min(BIC(spender))) ## and BIC instead
```

Extracting Coefficients Using Other Rules

```
## 1001 x 1 sparse Matrix of class "dgCMatrix"
##                                      seg31
## intercept                      5.929791309
## atdmt.com                      .
## yahoo.com                      .
## whenu.com                      .
```

```
## weatherbug.com                  .
## msn.com                         .
## google.com                      .
## aol.com                         .
## questionmarket.com              0.013756425
## googlesyndication.com-o02       .
## casalemedia.com                 .
## mywebsearch.com                 .
## myspace.com                     .
## pointroll.com                   .
## atwola.com                      .
## yieldmanager.com                .
## live.com                        .
## aim.com                         .
## mediaplex.com                   .
## precisionclick.com              .
## tribalfusion.com                .
## insightexpressai.com            .
## trafficmp.com                  -0.003495242
## ebay.com                        .
## realmedia.com                   .
## zedo.com                        .
## advertising.com                 .
## microsoft.com                   .
## hotbar.com                      .
## adrevolver.com                  .
## ru4.com                         .
## 180solutions.com                .
## nextag.com                      .
## accuweather.com                 .
## overture.com                    .
## hotmail.com                     .
## passport.com                    .
## my-etrust.com                   .
## starware.com                    .
## relevantknowledge.com          -0.034187615
## myway.com                       .
## partner2profit.com              .
## ditto.com                       .
## kanoodle.com                    .
## ebayobjects.com                 .
## mcafee.com                      .
## comcast.net                     0.007369149
## fastclick.net                   .
## adbrite.com                     .
## vpptechnologies.com             .
## specificclick.net               .
## serving-sys.com                 .
## weather.com                     .
## adserver.com                    .
## licenseacquisition.org          .
## pogo.com                        .
## go.com                          .
## btgrab.com                      .
```

```
## bellsouth.net                  .
## intellisrv.net                 .
## dell.com                       .
## waol.exe                       .
## cnn.com                        .
## facebook.com                   .
## incredibarvuz1.com             .
## burstnet.com                   .
## adknowledge.com                .
## funwebproducts.com             .
## belnk.com                      .
## netscape.com                   .
## mysearch.com                   .
## real.com                       .
## liveperson.net                 0.077066698
## adsonar.com                    .
## passport.net                   .
## euroclick.com                  .
## m7z.net                        .
## mywebface.com                  .
## kazaa.com                      .
## bestoffersnetworks.com         .
## vitalstream.com                .
## tacoda.net                     .
## unicast.com                    .
## offeroptimizer.com             .
## bankofamerica.com              .
## acsd.exe                       .
## gator.com                      .
## quickbrowsersearch.com         .
## revsci.net                     .
## personalweb.com                .
## rr.com                         .
## msnusers.com                   .
## zango.com                      .
## earthlink.net                  .
## mapquest.com                   .
## falkag.net                     .
## freeze.com                     .
## amazon.com                     0.054639845
## net-offers.net                 .
## shopperreports.com             .
## dellfix.com                    .
## plaxo.com                      .
## ysbweb.com                     .
## googleadservices.com           .
## qnsr.com                       .
## revenue.net                    .
## adultfriendfinder.com          .
## addynamix.com                  .
## seekmo.com                     .
## verizon.net                    .
## cox.net                        .
## metricsdirect.com              .
```

```
## akamai.net                      .
## admarketplace.net               .
## amazon.com-o01                  .
## aolacsd.exe                     .
## opinionsquare.com               .
## interclick.com                  .
## peoplepc.com                    .
## go.com-o04                      .
## realtechnetwork.net             .
## freezecoldcash.com              .
## ask.com                         .
## contextweb.com                  .
## intellitxt.com                  .
## yceml.net                       .
## about.com                       .
## youtube.com                     .
## wikipedia.org                   .
## surfaccuracy.com                .
## windowsmedia.com               -0.015026396
## craigslist.org                  .
## hackerwatch.org                 .
## foxsports.com                   .
## spamblockerutility.com          .
## walmart.com                     .
## navexcel.com                    .
## partypoker.com                  .
## wellsfargo.com                  .
## travelzoo.com                   .
## photobucket.com                 .
## viewpoint.com                   .
## nielsennetpanel.com             .
## mymailstamp.com                 .
## windows.com                     .
## optonline.net                   .
## eguard.com                      .
## aolcdn.com                      .
## musicmatch.com                  .
## qksz.net                        .
## cometsystems.com                .
## netzero.net                     .
## specificmedia.com               .
## paypal.com                      .
## iwon.com                        .
## monster.com-o01                 .
## vmn.net                         .
## juno.com                        .
## information.com                 .
## sysupdates.com                  .
## 2o7.net                         .
## adwave.com                      .
## need2find.com                   .
## target.com                      .
## ebayrtm.com                     0.092833283
## match.com                       .
```

```
## bridgetrack.com                .
## comcastsupport.com             .
## rs6.net                        .
## screensavers.com               .
## footprint.net                  .
## sportsline.com                 .
## adelphia.net                   .
## smileycentral.com              .
## dlqm.net                       .
## careerbuilder.com              .
## mlb.com                        .
## searchignite.com               .
## wachovia.com                   .
## expedia.com                    .
## thinktarget.com                .
## authnow.com                    .
## dotomi.com                     .
## blogspot.com                   .
## hpdjjs.com                     .
## chase.com                      .
## outerinfo.com                  .
## nscpcdn.com                    .
## vonage.com                     .
## searchscout.com                .
## compuserve.com                -0.004372746
## lycos.com                      .
## xanga.com                      .
## websearch.com                  .
## azjmp.com                      .
## tmcs.net-o01                   .
## exitexchange.com               .
## toshibapc.com                  .
## runescape.com                  .
## weatherstudio.com              .
## imdb.com                       .
## adecn.com                      .
## bargain-buddy.net              .
## carsdirect.com                 .
## mspaceads.com                  .
## apple.com                      .
## ups.com                        0.094620100
## 88.80.5.21                     .
## exct.net                       .
## cingular.com                   .
## foodnetwork.com                .
## go.com-o03                     .
## excite.com                     .
## capitalone.com                 .
## imiclk.com                     0.542497908
## overstock.com                  .
## bloglines.com                  .
## compfused.com                  .
## morpheus.com                   .
## foxnews.com                    .
```

```
## marketwatch.com                .
## wamu.com                       .
## monster.com                    .
## adobe.com                      .
## 888.com                        .
## untd.com                       .
## abetterinternet.com            .
## centralmedia.ws                .
## valuead.com                    .
## targetsaver.com                .
## lynxtrack.com                  .
## cartoonnetwork.com             .
## netflix.com                    .
## chitika.net                    .
## geocities.com                  .
## qsrch.com                      .
## drsnsrch.com                   .
## autobytel.com                  .
## web-nexus.net                  .
## webservicehosts.com            .
## sharewareonline.com            .
## llnwd.net                      .
## instantnavigation.com          .
## nick.com                       .
## nfl.com                        .
## oingo.com                      .
## lightningcast.net              .
## altbill.com                    .
## xolox.nl                       .
## superpages.com                 .
## classmates.com                 .
## aavalue.com                    .
## bluestreak.com                 .
## southwest.com                  0.019405476
## whitepages.com                 .
## usps.com                       .
## webhancer.com                  .
## bbc.co.uk                      .
## true.com                       .
## bearshare.com                  .
## citibank.com                   .
## blackplanet.com                .
## pch.com                        .
## att.net                        .
## autoweb.com                    .
## insightexpress.com             .
## charter.net                    .
## alumnigroup.org                .
## verizonwireless.com            .
## fedex.com                      .
## mobilesidewalk.com             .
## netteller.com                  .
## webshots.com                   .
## sprint.com                     .
```

```
## orbitz.com                      0.033293451
## bestbuy.com                     .
## grandstreetinteractive.com      .
## paypopup.com                    .
## cheaptickets.com                .
## dell4me.com                     .
## new.net                        -0.050232563
## nytimes.com                     .
## nyadmcncserve-05y06a.com        .
## aoltpspd.exe                    .
## toprebates.com                  .
## jcpenney.com                    0.289035701
## geotrust.com                    .
## travelocity.com                 0.241033879
## qvc.com                         0.116683951
## 4at1.com                        .
## cpmstar.com                     .
## bizrate.com                     .
## ticketmaster.com                .
## usbank.com                      .
## tripod.com                      .
## buy.com                         .
## nascar.com                      .
## aebn.net                        .
## infospace.com                   .
## wxbug.com                       .
## contextuads.com                 .
## bns1.net                        .
## download.com                    .
## gocyberlink.com                 .
## 192.168.1.1                     .
## dvlabs.com                      .
## defamer.com                     .
## tracking101.com                -0.217028216
## accountonline.com               .
## hbmediapro.com                  .
## usatoday.com                    .
## bigfishgames.com                .
## neopets.com                     .
## adoutput.com                    .
## sbc.com                         .
## noaa.gov                        .
## lowermybills.com                .
## kmpads.com                      .
## directtrack.com                 .
## clicksor.com                    .
## legacy.com                      .
## eajmp.com                       .
## nastydollars.com                .
## worldofwarcraft.com             .
## mirarsearch.com                 .
## verizon.com                     .
## miniclip.com                    .
## iwin.com                        .
```

```
## peel.com                        .
## hgtv.com                        .
## amaena.com                      .
## sprintpcs.com                   .
## shopping.com                    .
## webmd.com                       .
## clearchannel.com                .
## winamp.com                      .
## reference.com                   .
## interpolls.com                  .
## americangreetings.com           .
## tmcs.net                        0.427568831
## midtenmedia.com                 .
## domainsponsor.com               .
## thunderdownloads.com            .
## akamaistream.net                .
## livejournal.com                 .
## tx.us                           .
## onlinerewardcenter.com          .
## msn.com-o18                     .
## sony.com                        .
## dogpile.com                     .
## nba.com                         .
## citysearch.com                  0.148218793
## connextra.com                   .
## nickjr.com                      .
## t-mobile.com                    .
## winfixer.com                    .
## adlegend.com                    .
## adsrevenue.net                  .
## sears.com                       .
## ap.org                          .
## luna.net                        0.015598852
## shockwave.com                   .
## hsn.com                         .
## fl.us                           .
## mypoints.com                    .
## mozilla.org                     .
## aresgalaxy.org                  .
## realtor.com                     0.026481191
## addictinggames.com              .
## clickbooth.com                  .
## amateurmatch.com                .
## worldnow.com                    .
## surveys.com                     .
## pa.us                           .
## arcaderockstar.com              .
## coolsavings.com                -0.591097345
## yournewsletters.net             .
## liquidmedianetworks.com         .
## everythinggirl.com              .
## perfectmatch.com                .
## stockgroup.com                  .
## netster.com                     .
```

```
## bidclix.com                      .
## dropspam.com                     .
## hp.com                           .
## drivecleaner.com                 .
## consumerpromotioncenter.com      .
## aolwbspd.exe                     .
## americanexpress.com              0.019242435
## totaltalk.com                    .
## wwe.com                          .
## kontera.com                      .
## gamehouse.com                    .
## circuitcity.com                  .
## yimg.com                         .
## lightningcast.com                .
## edgefcs.net                      .
## wunderground.com                 .
## realarcade.com                   .
## singlesnet.com                   .
## azcentral.com                    .
## yellowpages.com                  .
## eharmony.com                     .
## paviliondownload.com             .
## insightbb.com                    .
## imageshack.us                    .
## shopzilla.com                    .
## ca.gov                           .
## donotchangeme.com                .
## ca.us                            .
## sourceforge.net                  .
## washingtonpost.com               .
## adjuggler.com                    .
## careercast.com                   .
## bangbros1.com                    .
## scripps.com-o01                  .
## migente.com                      .
## homedepot.com                    .
## winantivirus.com                 .
## irs.gov                          .
## blockbuster.com                  .
## kodakgallery.com                 .
## nih.gov                          .
## aol.com-o07                      .
## icq.com                          .
## wordcents.com                    .
## drudgereport.com                 .
## quizilla.com                     .
## srch-results.com                 .
## inqwire.com                      .
## ign.com                          .
## oinadserver.com                  .
## azoogleads.com                   .
## incredimail.com                  .
## shopathome.com                   .
## mtv.com                          .
```

```
## fidelity.com                    .
## bullseye-network.com            .
## flash-gear.com                  .
## proficient.com                  .
## autotrader.com                  .
## charter.com                     .
## healthology.com                 .
## evite.com                       .
## checkm8.com                    -0.024862981
## rsc01.net                       0.018608014
## oasei.com                       .
## heavy.com                       .
## slotch.com                      .
## passion.com                     .
## nbc.com                         .
## trafficmarketplace.com          .
## univision.com                   .
## priceline.com                   .
## flickr.com                      .
## andale.com                      .
## dealtime.com                    .
## yfdirect.com                    .
## entrepreneur.com                .
## go.com-o01                      .
## webmd.com-o01                   .
## sexsearch.com                   .
## pornaccess.com                  .
## gcion.com                       .
## shoplocal.com                   .
## kliptracker.com                 .
## nationalcity.com                .
## bbeplayer.com                   .
## videodome.com                   .
## 204.95.60.12                    .
## napster.com                     .
## myweather.net                   .
## msnbc.com                       .
## linkexchange.com                .
## searchmarketing.com             .
## angelfire.com                   .
## callwave.com                    .
## sonnerie.net                    .
## scout.com                       .
## rivals.com                      .
## altnet.com                      .
## spynet.com                      .
## macromedia.com                 -0.021666664
## ed.gov                          .
## wannawatch.com                  .
## frontiernet.net                 .
## flycell.com                     .
## edgesuite.net                   0.103259214
## 89.com                          .
## nc.us                           .
```

```
## ticketmaster.com-o01            .
## flowgo.com                      .
## cnet.com                        .
## oddcast.com                     .
## answers.com                     .
## timeinc.net                     .
## m5-systems.com                  .
## guideforyou.com                 .
## rn11.com                        .
## lowes.com                       .
## lifescript.com                  .
## shop.com                        .
## errorsafe.com                   .
## cams.com                        .
## macys.com                       0.305866313
## aa.com                          .
## addictingclips.com              .
## victoriassecret.com             0.760703451
## orchardbank.com                 .
## bravenet.com                    .
## imesh.com                       .
## nextel.com                      .
## screensandthemes.com            .
## suntrust.com                    .
## discovercard.com                .
## nbads.com                       .
## consumerincentiverewards.com    .
## valueclick.com                  .
## google.com-o03                  .
## cbs.com                         .
## bannerspace.com                 .
## technorati.com                  .
## cjt1.net                        .
## exactsearch.net                 .
## munky.com                       .
## cs.com                         -0.011421111
## kohls.com                       .
## tagged.com                      .
## babycenter.com                  .
## ebaumsworld.com                 .
## userplane.com                   .
## mediaplazza.com                 .
## netzerovoice.com                .
## gamespot.com                    .
## keen.com                        .
## bebo.com                        .
## rsc02.net                       0.171100930
## sysupdates2.com                 .
## imlive.com                      .
## oldnavy.com                     .
## regalinteractive.com            .
## weightwatchers.com              .
## subsag.com                      .
## aol.com-o08                     .
```

```
## azlyrics.com                    .
## freeringtonesnow.com            .
## freewebs.com                    .
## toysrus.com                     .
## hollywood.com                   .
## findwhat.com                    .
## local.com                       .
## webroot.com                     .
## tvguide.com                     .
## ny.us                           .
## resultsmaster.com               .
## jamster.com                     .
## gms1.net                        .
## switchboard.com                 .
## nicheseek.com                   .
## intelius.com                    .
## hi5.com                         .
## glispa.com                      .
## gannettonline.com               .
## cstv.com                        .
## adengage.com                    .
## superbrewards.com               .
## videocodezone.com               .
## symantecliveupdate.com          .
## pbskids.org                     .
## revresda.com                    .
## americansingles.com             .
## ugo.com-o02                     .
## job.com                         .
## installshield.com               .
## eprize.net                      .
## metacafe.com                    .
## focalex.com                     .
## cciads.us                       .
## perfectgonzo.com                .
## kbb.com                         .
## reunion.com                     .
## eproof.com                      .
## tripadvisor.com                 .
## bellsouth.com                   .
## search.com                      .
## comcast.com                     .
## ivillage.com                    .
## sun.com                         .
## regionsnet.com                  .
## mininova.org                    .
## beliefnet.com                   .
## intellicast.com                 .
## fastonlineusers.com             .
## gamespot.com-o01                .
## expedia.com-o01                 0.379543304
## military.com                    .
## musicnet.com                    .
## 53.com                          .
```

```
## oh.us                          .
## itrack.it                      .
## officedepot.com                0.274625984
## adultadworld.com               .
## univision.com-o01              .
## youravon.com                   .
## blackboard.com                 .
## yahoo.net                      0.228710425
## casinolasvegas.com             .
## warnerbros.com                 .
## delta.com                      0.036077437
## go.com-o02                     .
## deepnetexplorer.co.uk          .
## mozilla.com                    .
## opentracker.net                .
## break.com                      .
## catcha10.com                   .
## hotels.com                     .
## hallmark.com                   .
## sportsbook.com                 .
## mycheckfree.com                .
## ezboard.com                    .
## pro-market.net                 .
## mate1.com                      .
## awempire.com                   .
## jigzone.com                    .
## bangbrosnetwork.com            .
## marketlinx.com                 .
## tickle.com                    -0.319689884
## bbandt.com                     .
## mercuras.com                   .
## adtology2.com                  .
## bluemountain.com               .
## freepornofreeporn.com          .
## internet-optimizer.com         .
## autotrader.com-o01             .
## blogger.com                    .
## kraftfoods.com                 .
## loveaccess.com                 .
## shutterfly.com                 0.035946832
## stopzilla.com                  .
## xmradio.com                    .
## ga.us                          .
## ancestry.com                   .
## honda.com                      .
## fulltiltpoker.com              .
## il.us                          .
## ibsys.com                      .
## imixserver.com                 .
## barnesandnoble.com             .
## pricegrabber.com               .
## constantcontact.com            .
## zonelabs.com                   .
## pimpyourpro.com                .
```

```
## netflame.cc                     .
## slide.com                       .
## xnxx.com                        .
## upromise.com                    .
## livesexbar.com                  .
## videosz.com                     .
## freeweblayouts.net              .
## limewire.com                   -0.321858177
## ameritrade.com                  .
## freelaptop4you.com              .
## nickarcade.com                  .
## utkn.com                        .
## nj.us                           .
## 360i.com                        .
## finestresults.com               .
## asseenontvnetwork.com           .
## typepad.com                     .
## efax.com                        0.154058729
## regions.com                     .
## emachines.com                   .
## playaudiomessage.com            .
## bofunk.com                      .
## millsberry.com                  .
## cpvfeed.com                     .
## allrecipes.com                  .
## clubpenguin.com                 .
## eversave.com                    .
## ppmdating.com                   .
## lexico.com                      .
## usaa.com                        .
## directv.com                     .
## postini.com                     .
## secure-banking.com              .
## eyewonder.com                   .
## boston.com                      .
## ibanking-services.com           .
## astrology.com                   .
## datinggold.com                  .
## mlxchange.com                   .
## travelhook.net                  0.611641785
## custhelp.com                    .
## mn.us                           .
## zwire.com                       .
## emarketmakers.com               .
## gamefaqs.com                    .
## premiumproductsonline.com       .
## chrysler.com                    .
## prodigy.net                     .
## tv.com                          .
## windowsmedia.com-o04            .
## smashits.com                    .
## 65.115.67.11                    .
## snapfish.com                    .
## commerceonlinebanking.com       .
```

```
## bbt.com                          .
## linksynergy.com                  .
## yahoo.com-o08                    .
## freecodesource.com               .
## streamate.com                    .
## freecreditreport.com          -0.026140494
## intuit.com                       .
## rapid-pass.net                   .
## artistdirect.com                 .
## servedbyadbutler.com             .
## sidestep.com                     .
## adult.com                        .
## alltel.net                       .
## bcentral.com                     .
## openbank.com                     .
## nichedsites.com                  .
## cars.com                         .
## gm.com                           .
## adshuffle.com                    .
## freeslots.com                    .
## blink.com                        .
## candystand.com                   .
## monstermarketplace.com           .
## columbiahouse.com                0.084175381
## pncbank.com                      .
## discovery.com                    .
## hsbcbillpay.com                  .
## movietickets.com                 .
## page-not-found.net               .
## fandango.com                     .
## providianservices.com            .
## carad.com                        .
## homestead.com                    .
## realcastmedia.com                .
## webratsmusic.com                 .
## scottrade.com                    .
## cs102175.com                     .
## fnismls.com                      .
## shopperssavingcenter.com      -0.153796903
## hit-now.com                      .
## whatismyip.com                   .
## costco.com                       .
## bolt.com                         .
## bmgmusic.com                     .
## myhealthwealthandhappiness.com   .
## symantec.com                     .
## forbes.com                       .
## digitalcity.com                  .
## live365.com                      .
## firstadsolution.com              .
## linkconnector.com                .
## freepagegraphics.com             .
## imgfarm.com                      .
## insightexpresserdd.com           .
```

```
## pcsecurityshield.com           .
## allposters.com-o01             .
## msnvideo.com                   .
## miva.com                       .
## jackpotmadness.com             .
## mbnanetaccess.com              .
## newcarinsider.com              .
## edmunds.com                    .
## net-nucleus.com                .
## popcap.com                     .
## alt.com                        .
## staples.com                    0.769446720
## ussearch.com                   .
## bankone.com                    .
## rootv.com                      .
## citizensbankonline.com         .
## juggcrew.com                   .
## navyfcu.org                    .
## nordstrom.com                  0.330762740
## webstat.com                    .
## inklineglobal.com              .
## seeq.com                       .
## onetruemedia.com               .
## paltalk.com                    .
## sonypictures.com               .
## 204.181.57.155                 .
## commerceonline.com             .
## friendster.com                 .
## slate.com                      .
## hermoment.com                  .
## lovehappens.com                .
## mi.us                          .
## kmart.com                      .
## paidsurveys.com                .
## 123greetings.com               .
## blinko.com                     .
## citizensbank.com               .
## sirius.com                     .
## qrs1.net                       .
## adbureau.net                   .
## turn.com                       .
## abcdistributing.com            .
## fundsxpress.com                .
## pichunter.com                  .
## cbsnews.com                    .
## 216.139.222.230                .
## anywho.com                     .
## sedoparking.com                .
## householdbank.com              .
## treborwear.com                 .
## evault.ws                      .
## vh1.com                        .
## financialcontent.com           .
## gap.com                        0.053801713
```

```
## active.com                       .
## exclusivegiftcards.com           .
## michigan.gov                     .
## dada-mobile.net                  .
## textplussolutions.com            .
## myriadmarket.com                 .
## ifriends.net                     .
## aptimus.com                      .
## valueclick.net                   .
## pennyweb.com                     .
## blackpeoplemeet.com              .
## eltpath.com                      .
## yahoo.com-o46                    .
## sysprotect.com                   .
## dadamobile.com                   .
## cpxinteractive.com               .
## clickspring.net                  .
## staples-deals.com                .
## myyearbook.com                   .
## bravenetmedianetwork.com         .
## etrade.com                       .
## marykayintouch.com               .
## 64.39.16.166                     .
## moregamers.com                   .
## redorbit.com                     .
## tmz.com                          .
## blogrolling.com                  .
## checkfree.com                    .
## samsclub.com                     .
## va.us                            .
## united.com                       0.514951014
## certified-safe-downloads.com     .
## aimtoday.com                     .
## toseeka.com                      .
## bidz.com                         .
## gamespy.com                      .
## nylottery.org                    .
## godaddy.com                      .
## rsc03.net                        .
## altavista.com                    .
## ltdcommodities.com               .
## bhg.com                          .
## opm.gov                          .
## onlinemediaoutlet.com            .
## beboframe.com                    .
## cafepress.com                    .
## tarot.com                        .
## webgavel.com                     .
## rapmls.com                       .
## ztod.com                         .
## marriott.com                     0.256454941
## walgreens.com                    .
## rovion.com                       .
## ultimatebet.com                  .
```

```
## ea.com                          .
## petfinder.com                   .
## winsoftware.com                 .
## literotica.com                  .
## websourcedtraffic.com           .
## 032439.com                      .
## marketbanker.com                .
## clearchannelmusic.com           .
## colonize.com                    .
## searchfeed.com                  .
## eimg.net                        .
## shermanstravel.com              .
## key.com                         .
## multi-pops.com                  .
## yandex.ru                       .
## us.com                          .
## kinghost.com                    .
## sublimedirectory.com            .
## gogotools.com                   .
## camcrush.com                    .
## trafficexplorer.com             .
## myfamily.com                    .
## gay.com                         .
## freegiftworld.com               .
## dexonline.com                   .
## trade-in-value.com              .
## shopyourbargain.com             .
## dyndns.org                      .
## bizrate.com-o01                 0.453258899
## xctrk.com                       .
## webtoolcafe.com                 .
## zappos.com                      0.187201727
## wi.us                           .
## toptvbytes.com                  .
## 157.22.32.111                   .
## hotfreelayouts.com              .
## registrydefender.com            .
## zap2it.com                      .
## 64.136.28.49                   -0.038111674
## afy11.net                       .
## 207.97.212.250                  .
## invisionfree.com                .
## bravenet.com-o01                .
## gadgetcity.com                  .
## army.mil                        .
## yourgiftcards.com               .
## craigslist.com                  .
## usairways.com                   0.362587840
## drivelinemedia.com              .
## edline.net                      .
## dayport.com                     .
## axill.com                       .
## smartbargains.com               .
## newgrounds.com                  .
```

```
## 216.155.193.91                .
## providian.com                 .
## statcounter.com               .
## ajc.com                       .
## oprah.com                     .
## slingo.com                    .
## continental.com               0.234217772
## relevantchoice.com            .
## toontown.com                  .
## thumbplay.com                 .
## jacquielawson.com             .
## hotwire.com                   .
## nwa.com                       .
## atomz.com                     .
## nsgalleries.com               .
## uclick.com                    .
## mercurial.ca                  .
## schwab.com                    .
## nvero.net                     .
## ediets.com                    .
## ichotelsgroup.com             0.162595661
## 216.133.243.28                .
## aggregateknowledge.com        .
## topix.net                     .
## flalottery.com                .
## dlv4.com                      .
## mybloglog.com                 .
## lanxtra.com                   .
## away.com                      .
## grab.com                      .
## tipany.com                    .
## quickbooks.com                .
## instream.com                  .
## pbs.org                       .
## findology.com                 .
## business.com                  .
## cmt.com                       .
## myinsiderdeals.com            .
## imagine-msn.com               .
## nhl.com                       .
## modern-singles.net            .
## addfreestats.com              .
## rent.com                      .
## homegain.com                  .
## freeones.com                  .
## jetblue.com                   0.117016468
## loanweb.com                   .
## findarticles.com              .
## iwon.com-o04                  .
## incredigames.com              .
## webkinz.com                   .
## dealerconnection.com          .
## streamaudio.com               .
## grantmedia.com                .
```

```
## home123info.com                 .
## exittracking.com                .
## worldsex.com                    .
## yfdmedia.com                    .
## automotive.com                  .
## cursormania.com                 .
## tradedoubler.com                .
## bedbathandbeyond.com            0.093914776
## equifax.com                     .
## hotornot.com                    .
## falkag.de                       .
## chicagotribune.com              .
## airtran.com                     .
## thebreastcancersite.com         .
## charmingshoppes.com             .
## ugo.com                         .
## cox.com                         .
## spicymint.com                   .
## real.com-o01                    .
## targetnet.com                   .
## effectivebrand.com              .
## dallascowboys.com               .
## leadgenetwork.com               .
## in.us                           .
## vistaprint.com                  .
```

Cross-Validation

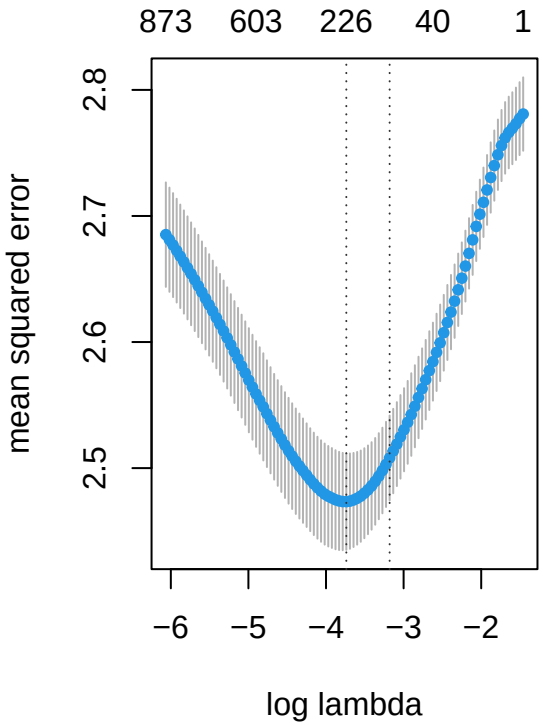
```
cv.spender <- cv.gamlr(xweb, log(yspend), verb=TRUE,nfold =10)
```

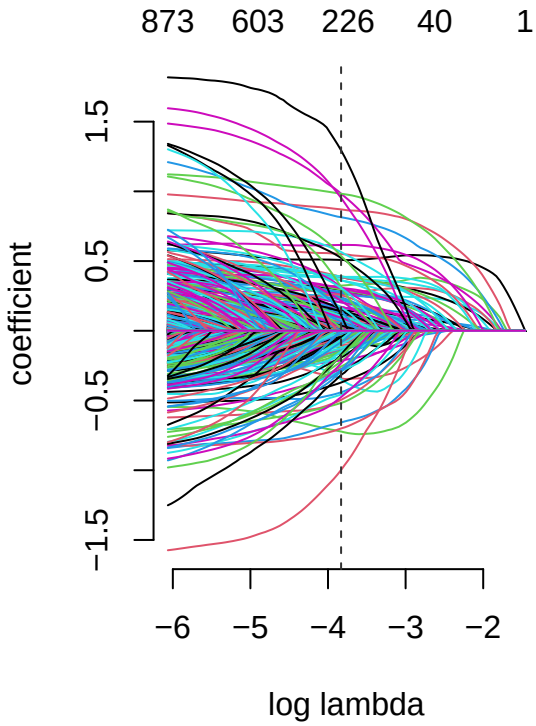
```
## fold 1,2,3,4,5,6,7,8,9,10,done.
```

```
beta1se <- coef(cv.spender) ## 1se rule; see ?cv.gamlr
betamin <- coef(cv.spender, select="min") ## min cv selection
cbind(beta1se,betamin)[c("tvguide.com","americanexpress.com"),]
```

```
## 2 x 2 sparse Matrix of class "dgCMatrix"
##                           seg38       seg50
## tvguide.com         .          -0.01207381
## americanexpress.com 0.04801744  0.04869937
```

```
par(mfrow=c(1,2))
plot(cv.spender)
plot(cv.spender$gamlr) ## cv.gamlr includes a gamlr object
```





Ploting CV Results

```
log(spender$lambda[which.min(AICc(spender))])
```

Log lambdas selected under various criteria

```
##     seg52
## -3.831328
```

```
log(spender$lambda[which.min(AIC(spender))])
```

```
##     seg62
## -4.296497
```

```
log(spender$lambda[which.min(BIC(spender))])
```

```
##     seg31
## -2.854474
```

```
log(cv.spender$lambda.min)
```

```
## [1] -3.738295
```

```
log(cv.spender$lambda.1se)
```

```
## [1] -3.180092
```

All Metrics, together in a path plot.

```
ll <- log(spender$lambda)

plot(spender, col="grey")
abline(v=ll[which.min(AICc(spender))], col="black", lty=2)
abline(v=ll[which.min(AIC(spender))], col="orange", lty=2)
abline(v=ll[which.min(BIC(spender))], col="green", lty=2)
abline(v=log(cv.spender$lambda.min), col="blue", lty=2)
abline(v=log(cv.spender$lambda.1se), col="purple", lty=2)
legend("topright", bty="n", lwd=1,
    col=c("black","orange","green","blue","purple"),
    legend=c("AICc","AIC","BIC","CV.min","CV.1se"))
```