SIS NMF surg rec

November 23, 2014

1 Preparation

```
options(java.parameters = "-Xmx4G")
library(survival)
## Loading required package: splines
library(energy)
library(NMF)
## Loading required package: methods
## Loading required package: pkgmaker
## Loading required package: registry
## Loading required package: rngtools
## Loading required package: cluster
## NMF - BioConductor layer [OK] | Shared memory capabilities [OK] | Cores 63/64
library(glmulti)
## Loading required package: rJava
## Attaching package: 'glmulti'
## The following object is masked from 'package:NMF':
##
##
     consensus
library(glmnet)
## Loading required package: Matrix
## Loaded glmnet 1.9-8
library(RColorBrewer)
library(gplots)
##
## Attaching package: 'qplots'
## The following object is masked from 'package:stats':
##
##
     lowess
library(xtable)
library(stargazer)
```

```
##
## Please cite as:
##
## Hlavac, Marek (2014). stargazer: LaTeX code and ASCII text for well-formatted regression
and summary statistics tables.
## R package version 5.1. http://CRAN.R-project.org/package=stargazer
setwd(data_dir)
load("image.rda")
```

2 Probe selection

```
table(cpss.sis$sel)

##
## FALSE TRUE
## 12805 195

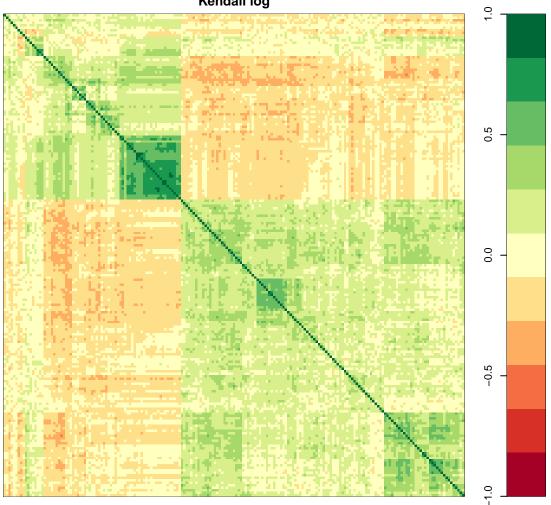
mean(cpss.sis$sel)

## [1] 0.015
```

3 Expression correlation

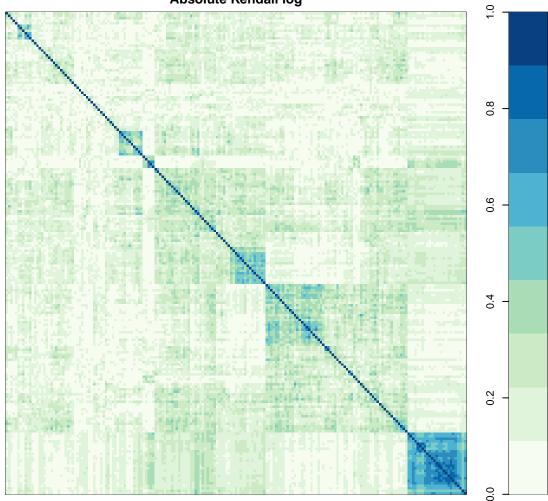
```
corPlot(x.sel.kcor, main = "Correlation Clusters of CPSS-SIS-FAST Probes\nKendall log",
    useRaster = FALSE)
```

Correlation Clusters of CPSS-SIS-FAST Probes Kendall log

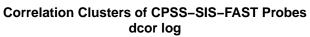


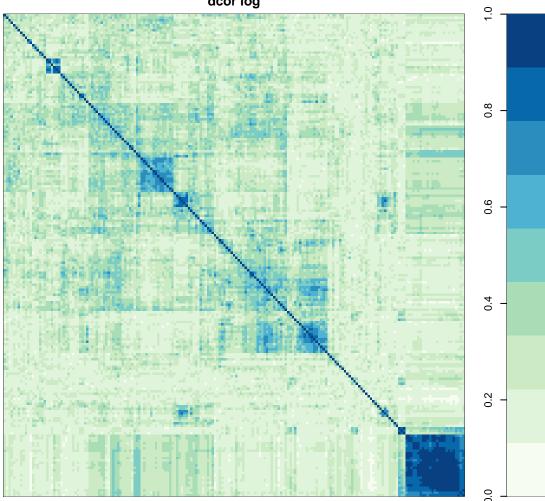
corPlot(abs(x.sel.kcor), zlim = c(0, 1), pal = "GnBu", main = "Correlation Clusters of CPSS-SIS-FAST Pro
 useRaster = FALSE)

Correlation Clusters of CPSS-SIS-FAST Probes Absolute Kendall log



corPlot(x.sel.dcor, zlim = c(0, 1), pal = "GnBu", main = "Correlation Clusters of CPSS-SIS-FAST Probes\nuseRaster = FALSE)

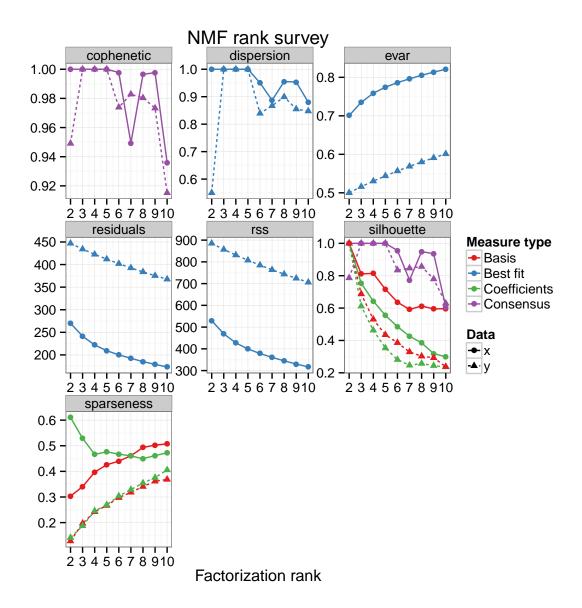




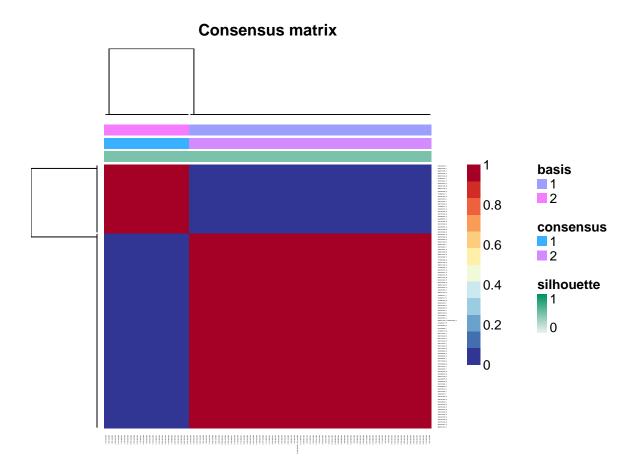
4 Factorization

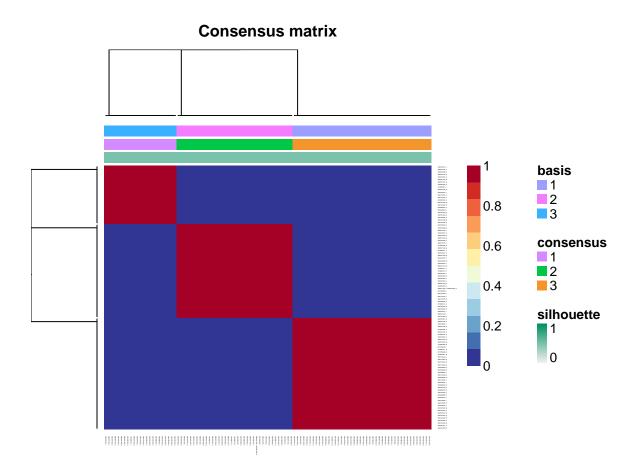
4.1 Rank estimation

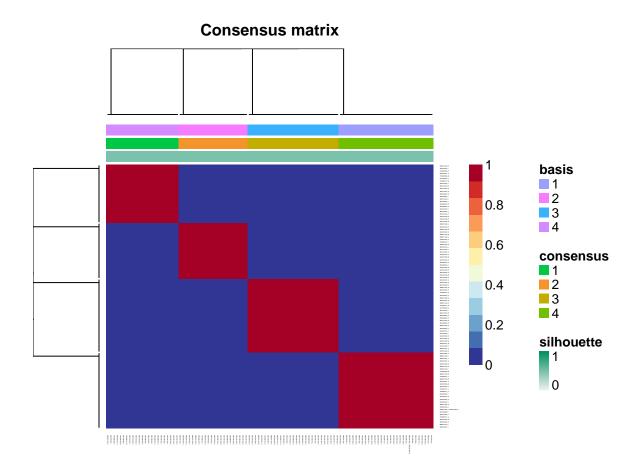
plot(temp.nmf.rank, temp.nmf.rank.random[[1]])

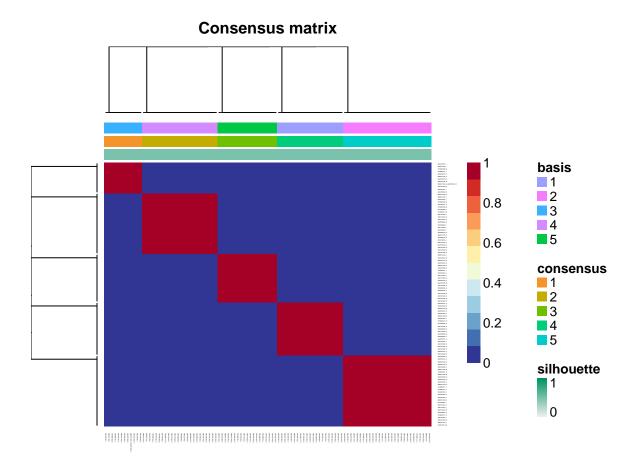


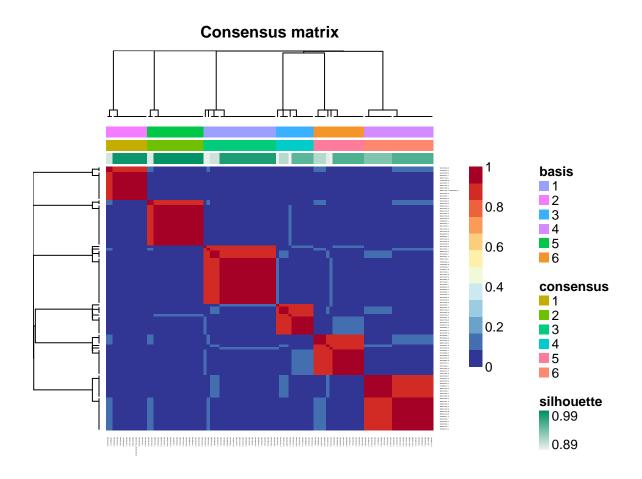
```
for (i in 1:length(temp.nmf.rank$fit)) {
    consensusmap(temp.nmf.rank$fit[[i]])
}
```

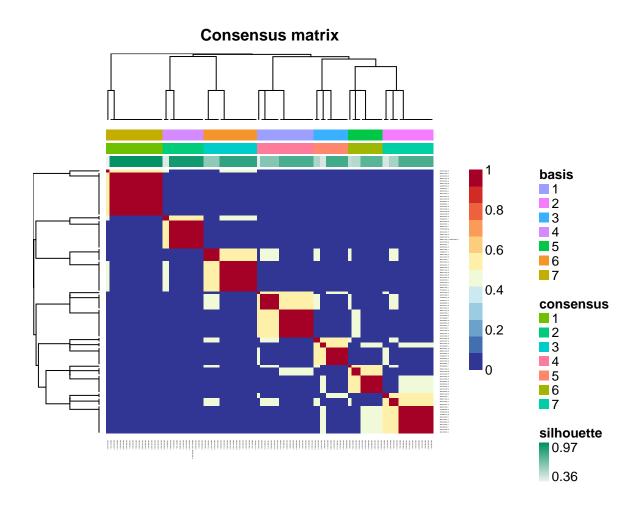


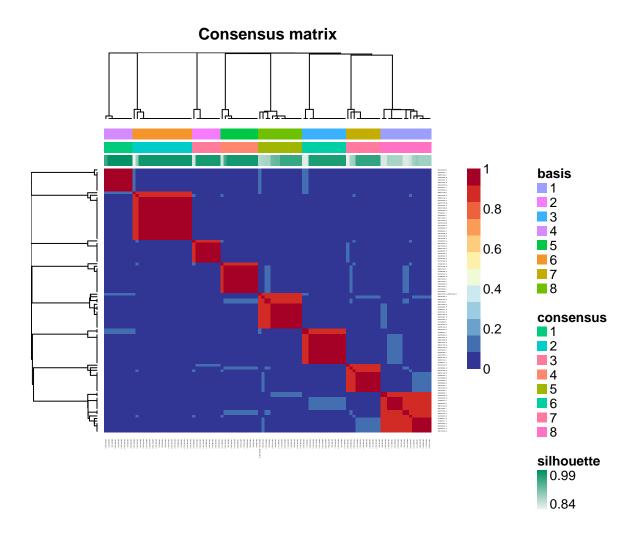


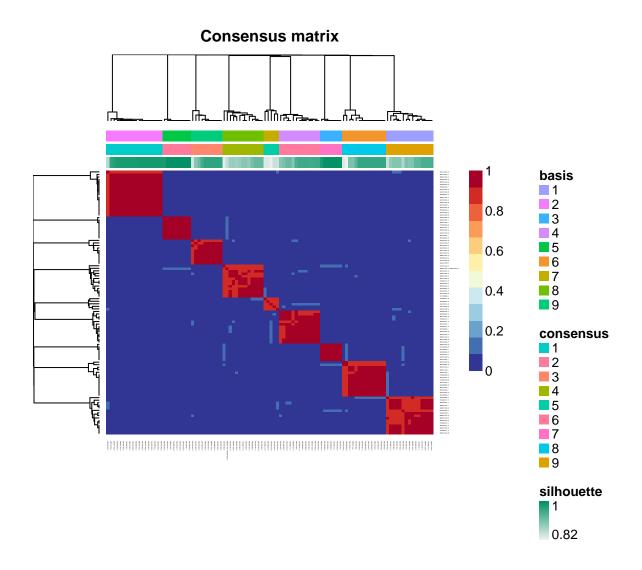


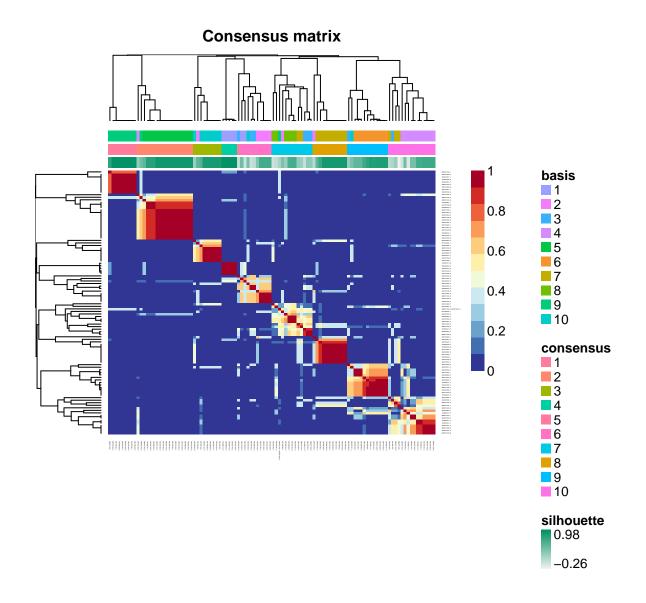






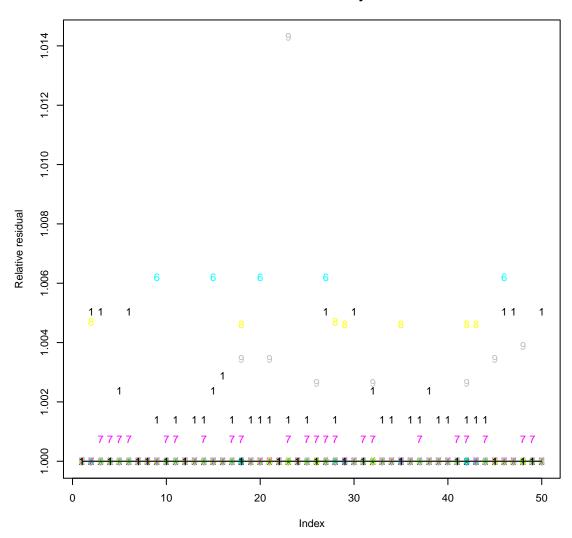






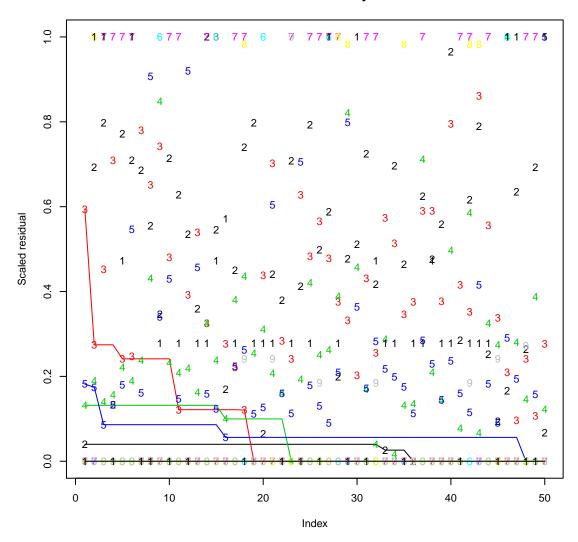
```
plot(0 ~ 0, type = "n", xlim = c(1, nrow(temp.resids)), ylim = range(temp.resids_rel),
    ylab = "Relative residual", main = "Solution Stability")
for (i in 1:ncol(temp.resids)) {
    points(temp.resids_rel[, i], col = i, pch = colnames(temp.resids)[i])
    lines(cummin(temp.resids_rel[, i]), col = i)
}
```

Solution Stability

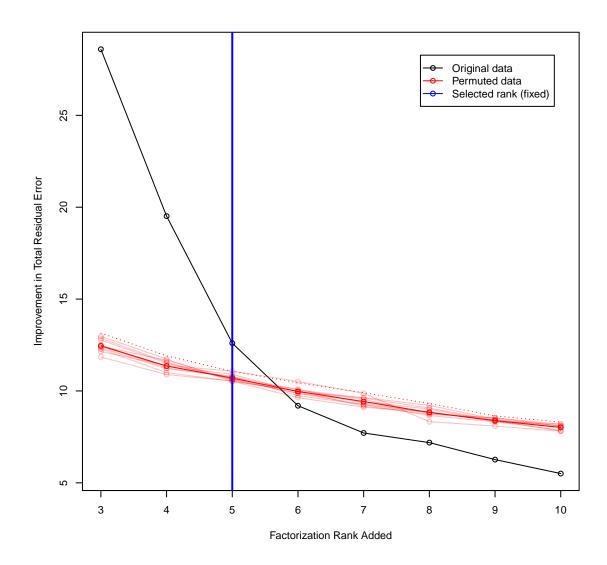


```
plot(0 ~ 0, type = "n", xlim = c(1, nrow(temp.resids)), ylim = range(temp.resids_scaled),
    ylab = "Scaled residual", main = "Solution Stability")
for (i in 1:ncol(temp.resids)) {
    points(temp.resids_scaled[, i], col = i, pch = colnames(temp.resids)[i])
    lines(cummin(temp.resids_scaled[, i]), col = i)
}
```

Solution Stability

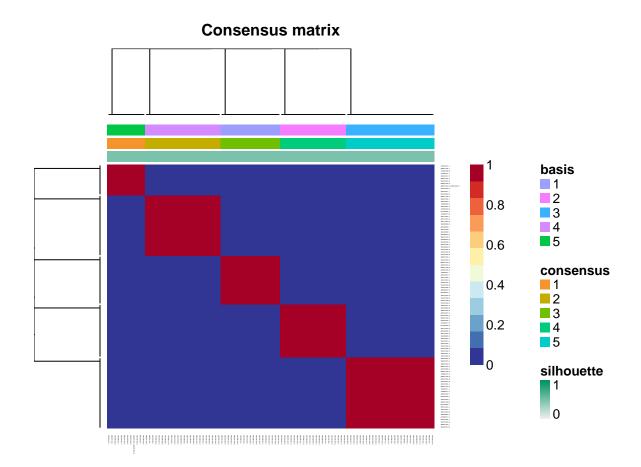


```
legend("topright", legend = c("Original data", "Permuted data", sprintf("Selected rank (%s)",
    ifelse(temp.col == "green", "auto", "fixed"))), col = c("black", "red",
    temp.col), lty = "solid", pch = 21, inset = 0.05)
```



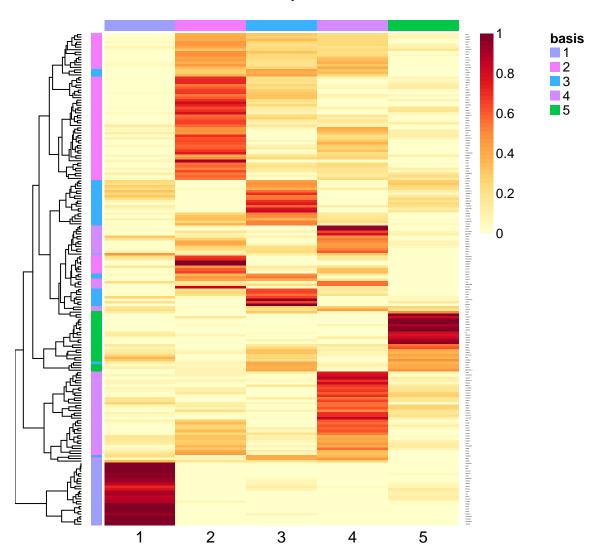
4.2 Fit

consensusmap(xlin.scaled.sel.nmf)

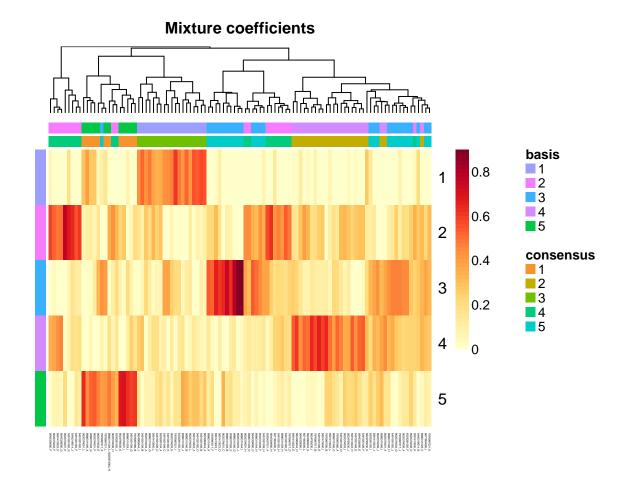


basismap(xlin.scaled.sel.nmf)

Basis components

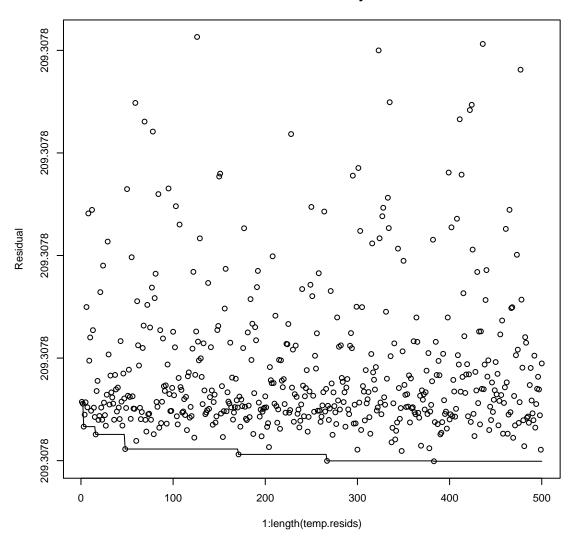


coefmap(xlin.scaled.sel.nmf)



```
temp.resids = sapply(xlin.scaled.sel.nmf, residuals)
plot(1:length(temp.resids), temp.resids, ylab = "Residual", main = "Solution Stability")
lines(1:length(temp.resids), cummin(temp.resids))
```

Solution Stability



4.3 Component CPV associations

4.3.1 Survival: Diagnosis to disease-specific death

```
for (i in 1:nrow(coef(xlin.scaled.sel.nmf))) {
    print(summary(coxph(y ~ coef(xlin.scaled.sel.nmf)[i, ])))
## Call:
   coxph(formula = y ~ coef(xlin.scaled.sel.nmf)[i, ])
##
##
     n= 104, number of events= 77
##
##
                                                                  z Pr(>|z|)
##
                                     coef exp(coef) se(coef)
   coef(xlin.scaled.sel.nmf)[i, ] -1.224
                                                       1.045 -1.17
##
                                   exp(coef) exp(-coef) lower .95 upper .95
##
```

Table 1:

_	$Dependent\ variable.$
	У
coef(xlin.scaled.sel.nmf)[i,]	-1.224
,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	(1.045)
Observations	104
\mathbb{R}^2	0.014
Max. Possible R ²	0.997
Log Likelihood	-307.100
Wald Test	1.370 (df = 1)
LR Test	1.473 (df = 1)
Score (Logrank) Test	1.383 (df = 1)
A.T. ,	* .0.1 ** .0.0* ***

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 2:

	Dependent variable:	
	У	
coef(xlin.scaled.sel.nmf)[i,]	7.725***	
, , , , , , , , , , , , , , , , , , , ,	(1.319)	
Observations	104	
\mathbb{R}^2	0.252	
Max. Possible R ²	0.997	
Log Likelihood	-292.800	
Wald Test	$34.310^{***} (df = 1)$	
LR Test		
Score (Logrank) Test	$36.300^{***} (df = 1)$	
Note:	*p<0.1; **p<0.05; ***p<	

Table 3:

	Dependent variable:
	У
coef(xlin.scaled.sel.nmf)[i,]	-1.441
, , , , , , , , , , , , , , , , , , , ,	(1.288)
Observations	104
\mathbb{R}^2	0.012
Max. Possible R ²	0.997
Log Likelihood	-307.200
Wald Test	1.250 (df = 1)
LR Test	1.299 (df = 1)
Score (Logrank) Test	1.255 (df = 1)
37 .	* 04 ** 00 ***

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 4:

	Dependent variable:
	у
coef(xlin.scaled.sel.nmf)[i,]	4.362***
, , , , , , , , , , , , , , , , , , , ,	(1.177)
Observations	104
\mathbb{R}^2	0.115
Max. Possible R ²	0.997
Log Likelihood	-301.500
Wald Test	$13.730^{***} (df = 1)$
LR Test	$12.660^{***} (df = 1)$
Score (Logrank) Test	$14.210^{***} (df = 1)$
Note:	*p<0.1; **p<0.05; ***p<

Table 5:

Table 9.	
	Dependent variable:
	У
coef(xlin.scaled.sel.nmf)[i,]	-6.189^{***}
· · · · · · · · · · · · · · · · · · ·	(1.687)
Observations	104
\mathbb{R}^2	0.148
Max. Possible \mathbb{R}^2	0.997
Log Likelihood	-299.500
Wald Test	$13.460^{***} (df = 1)$
LR Test	$16.640^{***} (df = 1)$
Score (Logrank) Test	$14.060^{***} (df = 1)$

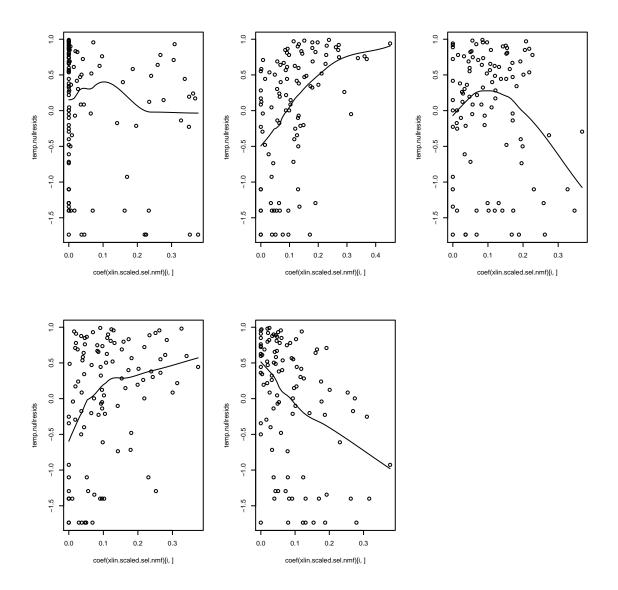
Note:

*p<0.1; **p<0.05; ***p<0.01

```
## coef(xlin.scaled.sel.nmf)[i, ] 0.294 3.4 0.0379 2.28
##
## Concordance= 0.546 (se = 0.035)
## Rsquare= 0.014 (max possible= 0.997)
## Likelihood ratio test= 1.47 on 1 df, p=0.225
## Wald test = 1.37 on 1 df,
                                      p=0.241
## Score (logrank) test = 1.38 on 1 df, p=0.24
## Call:
## coxph(formula = y ~ coef(xlin.scaled.sel.nmf)[i, ])
## n= 104, number of events= 77
##
                                coef exp(coef) se(coef) z Pr(>|z|)
##
## coef(xlin.scaled.sel.nmf)[i, ] 7.73 2264.99 1.32 5.86 4.7e-09
##
                             exp(coef) exp(-coef) lower .95 upper .95
## coef(xlin.scaled.sel.nmf)[i, ] 2265 0.000442 171 30045
##
## Concordance= 0.692 (se = 0.036)
## Rsquare= 0.252 (max possible= 0.997)
## Likelihood ratio test= 30.1 on 1 df, p=4.02e-08
## Wald test = 34.3 on 1 df, p=4.71e-09
## Score (logrank) test = 36.3 on 1 df, p=1.69e-09
## Call:
## coxph(formula = y ~ coef(xlin.scaled.sel.nmf)[i, ])
##
   n= 104, number of events= 77
##
                                coef exp(coef) se(coef) z Pr(>|z|)
## coef(xlin.scaled.sel.nmf)[i, ] -1.441 0.237 1.288 -1.12 0.26
##
##
                             exp(coef) exp(-coef) lower .95 upper .95
## coef(xlin.scaled.sel.nmf)[i, ] 0.237 4.23 0.0189 2.96
##
## Concordance= 0.508 (se = 0.036)
## Rsquare= 0.012 (max possible= 0.997)
## Likelihood ratio test= 1.3 on 1 df, p=0.254
## Wald test = 1.25 on 1 df,
                                     p=0.263
## Score (logrank) test = 1.26 on 1 df, p=0.263
##
## coxph(formula = y ~ coef(xlin.scaled.sel.nmf)[i, ])
##
## n= 104, number of events= 77
##
                              coef exp(coef) se(coef) z Pr(>|z|)
##
## coef(xlin.scaled.sel.nmf)[i, ] 4.36 78.42 1.18 3.71 0.00021
                              exp(coef) exp(-coef) lower .95 upper .95
##
## coef(xlin.scaled.sel.nmf)[i, ] 78.4 0.0128 7.8
```

```
## Concordance= 0.616 (se = 0.036)
## Rsquare= 0.115 (max possible= 0.997)
## Likelihood ratio test= 12.7 on 1 df, p=0.000374
## Wald test = 13.7 on 1 df, p=0.000211
## Score (logrank) test = 14.2 on 1 df, p=0.000163
## Call:
## coxph(formula = y ~ coef(xlin.scaled.sel.nmf)[i, ])
## n= 104, number of events= 77
##
                                   coef exp(coef) se(coef) z Pr(>|z|)
## coef(xlin.scaled.sel.nmf)[i, ] -6.18940 0.00205 1.68734 -3.67 0.00024
##
##
                               exp(coef) exp(-coef) lower .95 upper .95
## coef(xlin.scaled.sel.nmf)[i, ] 0.00205 488 7.51e-05 0.056
## Concordance= 0.649 (se = 0.036)
## Rsquare= 0.148 (max possible= 0.997)
## Likelihood ratio test= 16.6 on 1 df, p=4.52e-05
## Wald test = 13.5 on 1 df, p=0.000244
## Score (logrank) test = 14.1 on 1 df, p=0.000177
temp.nullfit = coxph(y ~ 1)
temp.nullresids = residuals(temp.nullfit, type = "martingale")
par(mfrow = c(2, 3))
for (i in 1:nrow(coef(xlin.scaled.sel.nmf))) {
   scatter.smooth(temp.nullresids ~ coef(xlin.scaled.sel.nmf)[i, ])
```

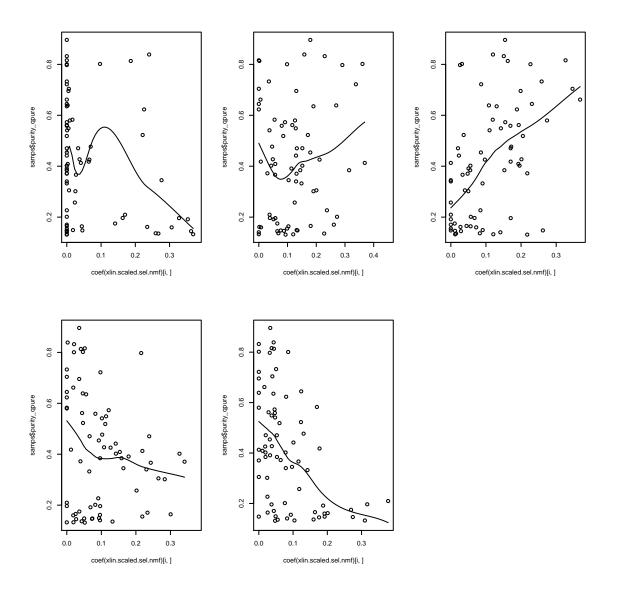
par(mfrow = c(1, 1))



4.4 Purity

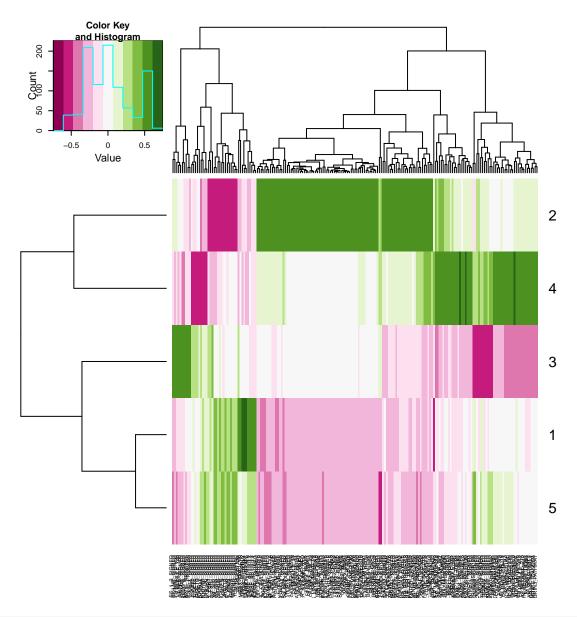
```
apply(coef(xlin.scaled.sel.nmf), 1, function(xc) cor.test(samps$purity_qpure,
    xc, method = "kendall"))
##
   [[1]]
##
##
   Kendall's rank correlation tau
##
## data: samps$purity_qpure and xc
## z = -1.903, p-value = 0.0571
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##
      tau
## -0.162
##
```

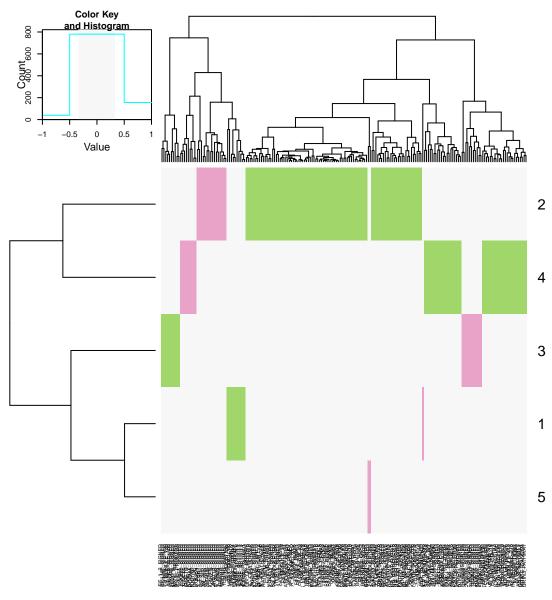
```
##
## [[2]]
## Kendall's rank correlation tau
##
## data: samps$purity_qpure and xc
## z = 0.5649, p-value = 0.5722
\#\# alternative hypothesis: true tau is not equal to 0
## sample estimates:
##
      tau
## 0.04493
##
##
## [[3]]
##
## Kendall's rank correlation tau
## data: samps$purity_qpure and xc
## z = 4.12, p-value = 3.785e-05
\#\# alternative hypothesis: true tau is not equal to 0
## sample estimates:
    tau
##
## 0.3288
##
##
## [[4]]
##
## Kendall's rank correlation tau
##
## data: samps$purity_qpure and xc
## z = -1.878, p-value = 0.06039
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##
      tau
## -0.1499
##
##
## [[5]]
##
## Kendall's rank correlation tau
## data: samps$purity_qpure and xc
## z = -3.865, p-value = 0.0001112
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##
      tau
## -0.3088
par(mfrow = c(2, 3))
for (i in 1:nrow(coef(xlin.scaled.sel.nmf))) {
    scatter.smooth(samps$purity_qpure ~ coef(xlin.scaled.sel.nmf)[i, ])
par(mfrow = c(1, 1))
```



4.5 MTC P-values

4.6 MSigDB score correlation thresholding





```
temp.sig_id = colnames(xlin.scaled.sel.nmf.msigdb.corr)
temp.sig_class = gsub("\\..*", "", temp.sig_id)
temp.nsigs = length(temp.sig_id)
temp.nmeta = nrow(xlin.scaled.sel.nmf.msigdb.corr)
tables = lapply(1:temp.nmeta, function(metagene_i) {
    tapply(1:temp.nsigs, temp.sig_class, function(sig_class_is) {
        all_cors = xlin.scaled.sel.nmf.msigdb.corr[, sig_class_is]
        this_cors = all_cors[metagene_i, ]
        this_ids = temp.sig_id[sig_class_is]

        all_sig_cors = abs(all_cors) >= sig.corr.threshold
        this_sig_cors = all_sig_cors[metagene_i, ]

        sigs_to_report = which(this_sig_cors)

if (length(sigs_to_report) == 0) {
        table = data.frame(GeneSet = c(), Correlation = c(), Metagenes = c())
```

```
} else {
            table = data.frame(GeneSet = this_ids[sigs_to_report], Correlation = this_cors[sigs_to_report]
                Metagenes = apply(all_cors[, sigs_to_report, drop = FALSE],
                  2, function(cors) {
                    sel = abs(cors) >= sig.corr.threshold
                    # A positive number implies that positive GSVA signal is associated with
                    # worse prognosis
                    paste(which(sel) * sign(cors[which(sel)]) * sign(xlin.scaled.sel.nmf.cpv.pvals$d.sur
                      collapse = ",")
                  }))
            table = table[order(-(table$Correlation)), ]
            rownames(table) <- NULL</pre>
        table
    }, simplify = FALSE)
})
tables
## [[1]]
## [[1]]$c1
## data frame with 0 columns and 0 rows
## [[1]]$c2
##
                                           GeneSet Correlation Metagenes
                            c2.BERGER_MBD2_TARGETS
                                                      0.6714
## 2 c2.REACTOME_DIGESTION_OF_DIETARY_CARBOHYDRATE
                                                       0.5204
                                                                      -1
                   c2.IGLESIAS_E2F_TARGETS_SIGNED
                                                       -0.5144
                                                                       1
##
## [[1]]$c3
          GeneSet Correlation Metagenes
## 1 c3.V$HNF1_Q6 0.5144
##
## [[1]]$c4
             GeneSet Correlation Metagenes
## 1 c4.GNF2_SPINK1
                         0.7429
                                         -1
## 2 c4.GNF2_SERPINI2
                           0.7223
                                         -1
##
## [[1]]$c5
##
                                   GeneSet Correlation Metagenes
## 1
             c5.CARBOXYPEPTIDASE_ACTIVITY
                                               0.5204
                                                              -1
## 2
        c5.SERINE_TYPE_PEPTIDASE_ACTIVITY
                                                0.5191
                                                              -1
## 3 c5.SERINE_TYPE_ENDOPEPTIDASE_ACTIVITY
                                                0.5153
                                                              -1
        c5.SERINE_HYDROLASE_ACTIVITY
                                                0.5140
                                                              -1
## 5
                              c5.DIGESTION
                                                0.5093
                                                              -1
##
## [[1]]$c6
## data frame with 0 columns and 0 rows
## [[1]]$c7
## data frame with 0 columns and 0 rows
##
## [[2]]
## [[2]]$c1
```

	dat	a frame with 0 columns and 0 rows	
##	ΓΓ2	·]]\$c2	
##		GeneSet	
##	1	c2.FOURNIER_ACINAR_DEVELOPMENT_LATE_2	
##	2	c2.WHITEFORD_PEDIATRIC_CANCER_MARKERS	
##	3	c2.SHEDDEN_LUNG_CANCER_POOR_SURVIVAL_A6	
##		c2.BASAKI_YBX1_TARGETS_SIGNED	
##		c2.WU_APOPTOSIS_BY_CDKN1A_VIA_TP53	
##		c2.GAVIN_FOXP3_TARGETS_CLUSTER_P6	
##		c2.CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_SIGNED	
##		c2.RHODES_UNDIFFERENTIATED_CANCER	
##		c2.SOTIRIOU_BREAST_CANCER_GRADE_1_VS_3_SIGNED	
##		c2.REACTOME_KINESINS	
##		c2.ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER	
##		c2.SMID_BREAST_CANCER_BASAL_SIGNED	
##		c2.BURTON_ADIPOGENESIS_PEAK_AT_24HR	
##		c2.OHASHI_AURKB_TARGETS c2.LOPEZ_MESOTELIOMA_SURVIVAL_TIME_SIGNED	
##		c2.EGFEZ_MESGTELTOMA_SGRVIVAL_TIME_STGNED	
##		c2.ZHOU_CELL_CYCLE_GENES_IN_IR_RESPONSE_24HR	
##		c2.RODRIGUES_THYROID_CARCINOMA_ANAPLASTIC_SIGNED	
##		c2.HOFFMANN_LARGE_TO_SMALL_PRE_BII_LYMPHOCYTE_SIGNED	
##		c2.WANG_METASTASIS_OF_BREAST_CANCER_ESR1_SIGNED	
##		c2.FARMER_BREAST_CANCER_CLUSTER_2	
##	22	c2.FINETTI_BREAST_CANCER_KINOME_RED	
##	23	c2.ZHAN_MULTIPLE_MYELOMA_PR_SIGNED	
##	24	c2.NAKAYAMA_SOFT_TISSUE_TUMORS_PCA2_SIGNED	
##	25	c2.REACTOME_PYRIMIDINE_METABOLISM	
##	26	c2.AMUNDSON_GAMMA_RADIATION_RESPONSE	
##	27	c2.BURTON_ADIPOGENESIS_3	
##	28	c2.REN_BOUND_BY_E2F	
##	29	c2.CHANG_CYCLING_GENES	
##		c2.CHANG_CORE_SERUM_RESPONSE_SIGNED	
##		c2.PID_FOXM1PATHWAY	
##		c2.MONTERO_THYROID_CANCER_POOR_SURVIVAL_SIGNED	
##		c2.WHITFIELD_CELL_CYCLE_LITERATURE	
##		c2.REACTOME_CELL_CYCLE_MITOTIC	
##		c2.REICHERT_MITOSIS_LIN9_TARGETS	
##		c2.EGUCHI_CELL_CYCLE_RB1_TARGETS	
##		c2.HONMA_DOCETAXEL_RESISTANCE	
##		c2.BOYAULT_LIVER_CANCER_SUBCLASS_G23_SIGNED c2.RODRIGUES_THYROID_CARCINOMA_POORLY_DIFFERENTIATED_SIGNED	
		c2.RODRIGUES_INTRUID_CARCINOMA_POURLY_DIFFERENTIATED_SIGNED c2.REACTOME_CYCLIN_A_B1_ASSOCIATED_EVENTS_DURING_G2_M_TRANSITION	
##		c2. REACTOME_CTCLIN_A_BI_ASSOCIATED_EVENTS_DURING_G2_M_TRANSITION c2. HONRADO_BREAST_CANCER_BRCA1_VS_BRCA2	
##		c2. CHUNG_BLISTER_CYTOTOXICITY_SIGNED	
##		c2.LEI_MYB_TARGETS	
##		c2.REACTOME_UNWINDING_OF_DNA	
##		c2.CHEMNITZ_RESPONSE_TO_PROSTAGLANDIN_E2_SIGNED	
##		c2.KEGG_CELL_CYCLE	
##		c2.LI_LUNG_CANCER	
##	48	c2.VANTVEER_BREAST_CANCER_ESR1_SIGNED	
##	49	c2.DOANE_BREAST_CANCER_ESR1_SIGNED	
##	50	c2.BROWNE_HCMV_INFECTION_2HR_SIGNED	

```
## 51
                             c2.LI_WILMS_TUMOR_VS_FETAL_KIDNEY_1_SIGNED
## 52
                           c2.MOLENAAR_TARGETS_OF_CCND1_AND_CDK4_SIGNED
## 53
                                              c2.RUIZ_TNC_TARGETS_SIGNED
## 54
                                c2.KOBAYASHI_EGFR_SIGNALING_24HR_SIGNED
## 55
                            c2.VANTVEER_BREAST_CANCER_METASTASIS_SIGNED
## 56
                   c2.STEIN_ESRRA_TARGETS_RESPONSIVE_TO_ESTROGEN_SIGNED
## 57
                                 c2.TANG_SENESCENCE_TP53_TARGETS_SIGNED
##
      Correlation Metagenes
## 1
           0.6067
                           2
## 2
           0.5977
                           2
## 3
           0.5947
                           2
## 4
           0.5917
                           2
## 5
           0.5831
## 6
           0.5808
                           2
                           2
## 7
           0.5636
                           2
## 8
           0.5613
## 9
           0.5595
                           2
                           2
## 10
           0.5493
## 11
           0.5493
                           2
                           2
## 12
           0.5433
                           2
## 13
           0.5407
                           2
## 14
           0.5399
                           2
## 15
           0.5377
                           2
## 16
           0.5369
## 17
           0.5343
                           2
                           2
## 18
           0.5336
                           2
## 19
           0.5328
                           2
## 20
           0.5302
## 21
           0.5268
                           2
## 22
           0.5264
                           2
## 23
           0.5234
                           2
                           2
## 24
           0.5227
                           2
## 25
           0.5227
## 26
           0.5223
                           2
## 27
                           2
           0.5215
## 28
           0.5212
                           2
                           2
## 29
           0.5212
## 30
           0.5197
                           2
                           2
## 31
           0.5189
## 32
           0.5178
                           2
                           2
## 33
           0.5118
## 34
           0.5095
                           2
                           2
## 35
           0.5088
                           2
## 36
           0.5080
## 37
           0.5080
                           2
## 38
           0.5065
                           2
                           2
## 39
           0.5050
## 40
                           2
           0.5050
                           2
## 41
           0.5047
                           2
## 42
           0.5043
## 43
           0.5024
                           2
                           2
## 44
           0.5009
## 45
           0.5005
                           2
         0.5005
## 46
```

```
## 47 0.5001
                       2
## 48
         -0.5058
                        -2
                        -2
## 49
         -0.5084
                        -2
## 50
         -0.5167
                        -2
## 51
         -0.5234
## 52
         -0.5377
                        -2
## 53
         -0.5508
                        -2
## 54
         -0.5640
                        -2
                        -2
## 55
         -0.5745
                        -2
## 56
         -0.5842
## 57
         -0.6075
                        -2
##
## [[2]]$c3
## data frame with 0 columns and 0 rows
## [[2]]$c4
           GeneSet Correlation Metagenes
## 1
      c4.MODULE_54
                    0.5696
## 2 c4.GNF2_ESPL1
                        0.5677
                                       2
                                       2
## 3 c4.GNF2_CDC20
                       0.5606
                                       2
## 4 c4.GNF2_H2AFX
                       0.5467
                                       2
## 5 c4.GNF2_CKS1B
                      0.5456
                                       2
## 6 c4.GNF2_MKI67
                        0.5441
## 7 c4.GNF2_HMMR
                        0.5407
                                       2
## 8 c4.GNF2_CCNB2
                        0.5388
                                       2
                                       2
## 9 c4.MODULE_105
                        0.5384
                                       2
## 10 c4.MODULE_57
                        0.5377
## 11 c4.GNF2_CENPF
                       0.5366
                                       2
## 12 c4.GNF2_CCNA2
                       0.5358
                                       2
## 13
        c4.GNF2_TTK
                        0.5351
                                       2
                                       2
## 14 c4.MODULE_253
                       0.5336
## 15 c4.GNF2_CDC2
                       0.5328
                                       2
                                       2
## 16 c4.MODULE_126
                        0.5324
## 17 c4.MODULE 18
                        0.5290
                                       2
                                       2
## 18 c4.GNF2_RFC4
                        0.5230
## 19 c4.MODULE_17
                       0.5219
                                       2
## 20 c4.GNF2_BUB1
                                       2
                        0.5200
                                       2
## 21 c4.GNF2_MCM4
                        0.5148
                                       2
## 22 c4.MODULE_3
                        0.5148
## 23 c4.GNF2_BUB1B
                        0.5110
                                       2
                                       2
## 24 c4.GNF2_CENPE
                        0.5099
## 25 c4.GNF2_RFC3
                        0.5092
                                       2
                                       2
## 26 c4.GNF2_RRM2
                        0.5077
## 27 c4.GNF2_PCNA
                                       2
                        0.5065
## 28 c4.GNF2_CKS2
                        0.5047
                                       2
                                       2
## 29 c4.MODULE_315
                        0.5043
## 30 c4.MODULE_244
                        0.5032
##
## [[2]]$c5
##
                                            GeneSet Correlation Metagenes
## 1
                    c5.M_PHASE_OF_MITOTIC_CELL_CYCLE
                                                         0.5790
## 2
                                                         0.5771
                                                                        2
                           c5.CELL_CYCLE_GO_0007049
## 3
                                         c5.M_PHASE
                                                         0.5745
                                                                        2
## 4
                                         c5.MITOSIS 0.5733
```

```
## 5
                                c5.MITOTIC_CELL_CYCLE
                                                           0.5628
## 6
                             c5.REGULATION_OF_MITOSIS
                                                            0.5576
                                                                           2
## 7
                                                                           2
                                c5.CELL_CYCLE_PROCESS
                                                            0.5489
                                                                           2
## 8
                 c5.CELL_CYCLE_CHECKPOINT_GO_0000075
                                                            0.5471
                                                                           2
## 9
                                  c5.CELL_CYCLE_PHASE
                                                            0.5414
## 10
              c5.SPINDLE_ORGANIZATION_AND_BIOGENESIS
                                                                           2
                                                            0.5260
## 11
                         c5.REGULATION_OF_CELL_CYCLE
                                                            0.5230
                                                                           2
## 12
                    c5.MITOTIC_CELL_CYCLE_CHECKPOINT
                                                                           2
                                                           0.5215
                                                                           2
                     c5.NUCLEOTIDE_METABOLIC_PROCESS
                                                            0.5099
## 13
                                                                           2
## 14 c5.MITOTIC_SPINDLE_ORGANIZATION_AND_BIOGENESIS
                                                            0.5077
## [[2]]$c6
## data frame with 0 columns and 0 rows
##
## [[2]]$c7
##
                                                                GeneSet.
## 1 c7.GSE30962_PRIMARY_VS_SECONDARY_ACUTE_LCMV_INF_CD8_TCELL_SIGNED
             c7.GSE24634_NAIVE_CD4_TCELL_VS_DAY3_IL4_CONV_TREG_SIGNED
## 3
                       c7.GSE3982_CENT_MEMORY_CD4_TCELL_VS_TH2_SIGNED
## 4
       c7.GSE36476_CTRL_VS_TSST_ACT_72H_MEMORY_CD4_TCELL_YOUNG_SIGNED
## 5
                             c7.GSE3982_MEMORY_CD4_TCELL_VS_TH1_SIGNED
                                        c7.GSE3982_BCELL_VS_TH1_SIGNED
## 6
## 7
       c7.GSE36476_CTRL_VS_TSST_ACT_40H_MEMORY_CD4_TCELL_YOUNG_SIGNED
     Correlation Metagenes
         0.5715
## 1
                         2
                         -2
## 2
         -0.5009
                        -2
## 3
         -0.5043
## 4
         -0.5275
                        -2
## 5
         -0.5298
                        -2
## 6
         -0.5448
                        -2
## 7
         -0.5598
                        -2
##
## [[3]]
## [[3]]$c1
## data frame with 0 columns and 0 rows
##
## [[3]]$c2
##
                                                      GeneSet Correlation
## 1 c2.CHARAFE_BREAST_CANCER_LUMINAL_VS_MESENCHYMAL_SIGNED
                                                                    0.5843
## 2
                      c2.SMID_BREAST_CANCER_LUMINAL_B_SIGNED
                                                                    0.5677
## 3
        c2.CHARAFE_BREAST_CANCER_BASAL_VS_MESENCHYMAL_SIGNED
                                                                    0.5178
## 4
                                c2.LIU_PROSTATE_CANCER_SIGNED
                                                                    0.5129
## 5
                    c2.WAMUNYOKOLI_OVARIAN_CANCER_LMP_SIGNED
                                                                   0.5129
## 6
             c2.WAMUNYOKOLI_OVARIAN_CANCER_GRADES_1_2_SIGNED
                                                                    0.5072
## 7
                 c2.SENGUPTA_NASOPHARYNGEAL_CARCINOMA_SIGNED
                                                                   -0.5042
## 8
                       c2.LINDGREN_BLADDER_CANCER_CLUSTER_2B
                                                                   -0.5102
## 9
                       c2.WOO_LIVER_CANCER_RECURRENCE_SIGNED
                                                                   -0.5208
             c2.SCHUETZ_BREAST_CANCER_DUCTAL_INVASIVE_SIGNED
## 10
                                                                   -0.5373
## 11
       c2.LIEN_BREAST_CARCINOMA_METAPLASTIC_VS_DUCTAL_SIGNED
                                                                   -0.5493
## 12
                             c2.LIM MAMMARY STEM CELL SIGNED
                                                                   -0.5497
## 13
                             c2.ROY_WOUND_BLOOD_VESSEL_SIGNED
                                                                   -0.5580
## 14
           c2.VECCHI_GASTRIC_CANCER_ADVANCED_VS_EARLY_SIGNED
                                                                   -0.5689
                c2.SERVITJA_ISLET_HNF1A_TARGETS_SIGNED
## 15
                                                                  -0.5771
```

```
## Metagenes
## 1
       -3
## 2
            -3
## 3
            -3
## 4
           -3
## 5
            -3
## 6
            -3
## 7
            3
## 8
            3
## 9
            3
## 10
            3
## 11
             3
## 12
            3
## 13
            3
## 14
             3
## 15
             3
## [[3]]$c3
## data frame with 0 columns and 0 rows
## [[3]]$c4
## GeneSet Correlation Metagenes
## 1 c4.MODULE_180 0.5813 -3
## 2 c4.MODULE_139
                     0.5369
                                   -3
##
## [[3]]$c5
##
                                GeneSet Correlation Metagenes
## 1 c5.GLYCOPROTEIN_BIOSYNTHETIC_PROCESS 0.5238 -3
## 2 c5.GLYCOPROTEIN_METABOLIC_PROCESS
                                          0.5125
                                                         -3
## 3
      c5.PEPTIDYL_TYROSINE_MODIFICATION
                                        -0.5016
                                                         3
##
## [[3]]$c6
               GeneSet Correlation Metagenes
## 1 c6.LEF1_UP.V1_SIGNED -0.5937
##
## [[3]]$c7
## data frame with 0 columns and 0 rows
##
##
## [[4]]
## [[4]]$c1
## data frame with 0 columns and 0 rows
## [[4]]$c2
##
## 1
                                       c2.PID_INTEGRIN5_PATHWAY
## 2
                                       c2.PID_SYNDECAN_1_PATHWAY
## 3
                                       c2.BURTON_ADIPOGENESIS_8
         c2.ANASTASSIOU_CANCER_MESENCHYMAL_TRANSITION_SIGNATURE
## 5 c2.KEGG_ARRHYTHMOGENIC_RIGHT_VENTRICULAR_CARDIOMYOPATHY_ARVC
## 6
                                       c2.PID_INTEGRIN1_PATHWAY
## 7
                                        c2.PID_UPA_UPAR_PATHWAY
## 8
                                       c2.PID_INTEGRIN3_PATHWAY
## 9
                                  c2.POTTI_TOPOTECAN_SENSITIVITY
```

```
## 10
                                              c2.KEGG_FOCAL_ADHESION
                                      c2.REACTOME_COLLAGEN_FORMATION
## 11
                                   c2.KEGG_ECM_RECEPTOR_INTERACTION
## 12
                c2.REACTOME_CELL_EXTRACELLULAR_MATRIX_INTERACTIONS
## 13
               c2.YAO_TEMPORAL_RESPONSE_TO_PROGESTERONE_CLUSTER_16
## 14
## 15
                              c2.VERRECCHIA_EARLY_RESPONSE_TO_TGFB1
## 16
                c2.HELLEBREKERS_SILENCED_DURING_TUMOR_ANGIOGENESIS
## 17
                                        c2.PID_INTEGRIN_A9B1_PATHWAY
## 18
                      c2.REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION
## 19
                                          c2.PID_INTEGRIN_CS_PATHWAY
## 20
                           c2.KEGG_REGULATION_OF_ACTIN_CYTOSKELETON
## 21
                                                  c2.PID_FAK_PATHWAY
## 22
                                                c2.GILDEA_METASTASIS
## 23
                                        c2.PID_AVB3_INTEGRIN_PATHWAY
## 24
                                           c2.PID_SYNDECAN_4_PATHWAY
## 25
                                                c2.WU_CELL_MIGRATION
## 26
                                          c2.KARAKAS TGFB1 SIGNALING
## 27
                         c2.WIEDERSCHAIN_TARGETS_OF_BMI1_AND_PCGF2
## 28
                                      c2.AGARWAL_AKT_PATHWAY_TARGETS
## 29
                         c2.YAMASHITA_METHYLATED_IN_PROSTATE_CANCER
## 30
                                  c2.AMIT_SERUM_RESPONSE_240_MCF10A
## 31
                                      c2.KRIEG_HYPOXIA_NOT_VIA_KDM3A
## 32
                          c2.ELVIDGE_HIF1A_AND_HIF2A_TARGETS_SIGNED
## 33
                            c2.REN_ALVEOLAR_RHABDOMYOSARCOMA_SIGNED
## 34
                                      c2.PASINI_SUZ12_TARGETS_SIGNED
## 35
                    c2.MIYAGAWA_TARGETS_OF_EWSR1_ETS_FUSIONS_SIGNED
             c2.RICKMAN_TUMOR_DIFFERENTIATED_WELL_VS_POORLY_SIGNED
## 36
      Correlation Metagenes
## 1
           0.6182
                           4
## 2
           0.6088
                           4
## 3
                           4
           0.6080
           0.6002
                           4
                           4
## 5
           0.5930
## 6
           0.5852
                           4
## 7
           0.5642
                           4
## 8
           0.5600
                           4
## 9
                           4
           0.5597
                           4
## 10
           0.5481
                           4
## 11
           0.5481
## 12
           0.5473
                           4
## 13
           0.5451
                           4
## 14
           0.5357
                           4
## 15
           0.5278
                           4
## 16
           0.5244
                           4
## 17
           0.5233
                           4
## 18
                           4
           0.5196
## 19
           0.5192
                           4
## 20
           0.5177
                           4
## 21
           0.5132
                           4
## 22
           0.5124
                           4
## 23
           0.5106
## 24
           0.5106
                           4
## 25
           0.5076
                           4
## 26
           0.5061
```

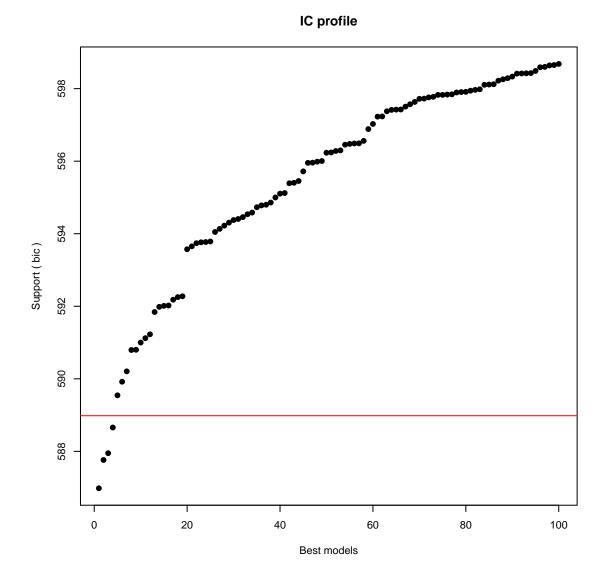
```
## 27 0.5034
## 28
          0.5019
                         4
                         4
## 29
          0.5016
                         4
## 30
          0.5004
## 31
                         4
         0.5004
## 32
        -0.5132
                        -4
## 33
        -0.5211
                        -4
## 34
        -0.5496
                        -4
## 35
         -0.5803
                        -4
## 36
         -0.5870
                        -4
##
## [[4]]$c3
## data frame with 0 columns and 0 rows
##
## [[4]]$c4
##
          GeneSet Correlation Metagenes
## 1 c4.GNF2_MMP1 0.5904
## 2 c4.MODULE_321
                       0.5694
                                     4
## 3 c4.MODULE_562
                      0.5526
                                     4
## 4 c4.GNF2_CDH11
                     0.5439
## 5 c4.GNF2_PTX3
                     0.5376
## 6 c4.MODULE_275
                                     4
                     0.5124
## 7 c4.MODULE_412
                     0.5019
                                     4
## [[4]]$c5
##
                  GeneSet Correlation Metagenes
## 1
             c5.COLLAGEN 0.5316 4
       c5.AXON_GUIDANCE
                              0.5237
## 3 c5.TISSUE_DEVELOPMENT
                             0.5233
                                             4
##
## [[4]]$c6
## data frame with 0 columns and 0 rows
## [[4]]$c7
                                                             GeneSet
##
                                       c7.GSE3982_DC_VS_BCELL_SIGNED
## 1
## 2
                                      c7.GSE3982_MAC_VS_BCELL_SIGNED
## 3
                                 c7.GSE27786_NKCELL_VS_NKTCELL_SIGNED
## 4
                                 c7.GSE22886_NAIVE_TCELL_VS_DC_SIGNED
## 5
                                c7.GSE29618_BCELL_VS_MONOCYTE_SIGNED
## 6
               c7.GSE1460_CD4_THYMOCYTE_VS_THYMIC_STROMAL_CELL_SIGNED
## 7 c7.GSE1460_INTRATHYMIC_T_PROGENITOR_VS_THYMIC_STROMAL_CELL_SIGNED
## Correlation Metagenes
## 1
        0.5417
## 2
         0.5083
                       4
## 3
                       4
        0.5016
## 4
        -0.5012
                       -4
## 5
        -0.5031
                       -4
## 6
        -0.5514
                       -4
## 7
        -0.6043
                       -4
##
##
## [[5]]
## [[5]]$c1
```

```
## [[5]]$c2
##
                            GeneSet Correlation Metagenes
## 1 c2.HAHTOLA_SEZARY_SYNDROM_SIGNED -0.5089
            c2.WINTER_HYPOXIA_SIGNED
                                        -0.5300
                                                       5
##
## [[5]]$c3
## data frame with 0 columns and 0 rows
## [[5]]$c4
## data frame with 0 columns and 0 rows
## [[5]]$c5
## data frame with 0 columns and 0 rows
##
## [[5]]$c6
## data frame with 0 columns and 0 rows
##
## [[5]]$c7
## data frame with 0 columns and 0 rows
print(asreg.result)
## glmulti.analysis
## Method: h / Fitting: coxph / IC used: bic
## Level: 2 / Marginality: TRUE
## From 100 models:
## Best IC: 586.982783576264
## Best model:
## [1] "Surv(time, event) ~ 1 + mg.2 + mg.4"
## Evidence weight: 0.178839068714947
## Worst IC: 598.680887369867
## 4 models within 2 IC units.
## 47 models to reach 95% of evidence weight.
coef(asreg.result)
            Estimate Uncond. variance Nb models Importance +/- (alpha=0.05)
## mg.1:mg.5 -0.04305
                       0.01590 3 0.004536
                                                                   0.2501
## mg.1:mg.3 0.06329
                            0.02841
                                           5 0.005286
                                                                   0.3344
## mg.1:mg.4 0.08138
                                           8 0.010086
                             0.06786
                                                                   0.5168
## mg.3:mg.5 0.13560
                             0.19571
                                            8 0.010353
                                                                   0.8776
## mg.1:mg.2 -0.32611
                                           12 0.019221
                            0.61494
                                                                   1.5556
## mg.4:mg.5 0.05730
                            0.30744
                                           6 0.021360
                                                                   1.0999
                                           13 0.046133
## mg.2:mg.5 -0.44153
                             2.68816
                                                                   3.2525
## mg.3:mg.4 1.29553
                             7.77189
                                           17 0.053343
                                                                  5.5303
## mg.2:mg.3 1.52913
                           11.06376
                                          20 0.062866
                                                                  6.5984
                                            25 0.143015
## mg.2:mg.4 -2.69795
                            29.94018
                                                                 10.8546
        0.12389
## mg.1
                             0.12171
                                            53 0.154331
                                                                   0.6921
## mg.3
           0.18971
                            1.30123
                                            66 0.330022
                                                                   2.2629
## mg.5
            -1.79547
                            5.65647
                                            58 0.467331
                                                                   4.7180
## mg.4
           2.72919
                            5.25524
                                           72 0.720460
                                                                   4.5476
                                         96 0.990789
## mg.2 7.27130
                           4.08075
                                                                   4.0073
```

data frame with 0 columns and 0 rows

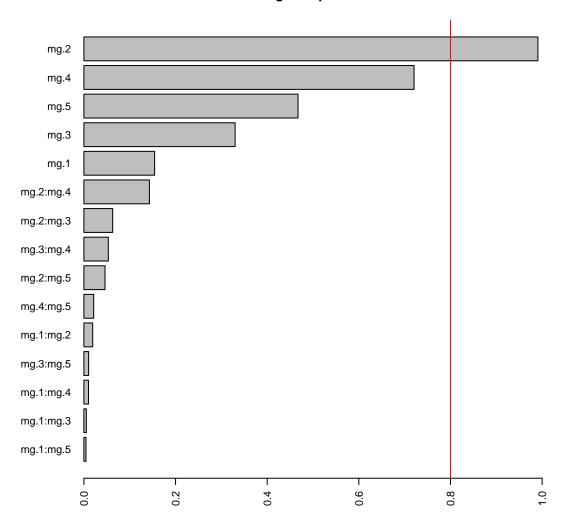
```
summary(asreg.result@objects[[1]])
## Call:
## fitfunc(formula = as.formula(x), data = data)
## n= 104, number of events= 77
##
        coef exp(coef) se(coef) z Pr(>|z|)
## mg.2 7.21 1350.26 1.35 5.33 9.6e-08
## mg.4 3.47 32.26 1.25 2.78 0.0054
##
## exp(coef) exp(-coef) lower .95 upper .95
## mg.2 1350.3 0.000741 95.53 19085
## mg.4 32.3 0.031001 2.79
                                    374
##
## Concordance= 0.706 (se = 0.036)
## Rsquare= 0.302 (max possible= 0.997)
## Likelihood ratio test= 37.4 on 2 df, p=7.48e-09
## Wald test = 40.3 on 2 df, p=1.81e-09
## Score (logrank) test = 42.5 on 2 df, p=5.8e-10
```

```
plot(asreg.result, type = "p")
```



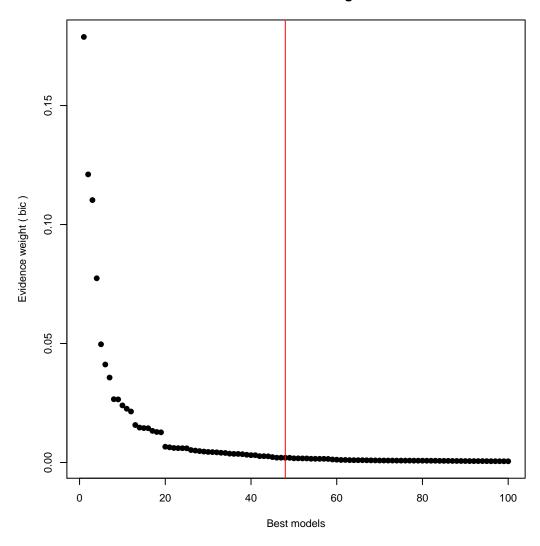
plot(asreg.result, type = "s")

Model-averaged importance of terms



plot(asreg.result, type = "w")

Profile of model weights



```
glmnet.coef.1se

## 5 x 1 sparse Matrix of class "dgCMatrix"

## mg.1 .

## mg.2 3.9181

## mg.3 .

## mg.5 -1.3025

glmnet.coef.min

## 5 x 1 sparse Matrix of class "dgCMatrix"

## mg.1 .

## mg.1 .

## mg.1 .

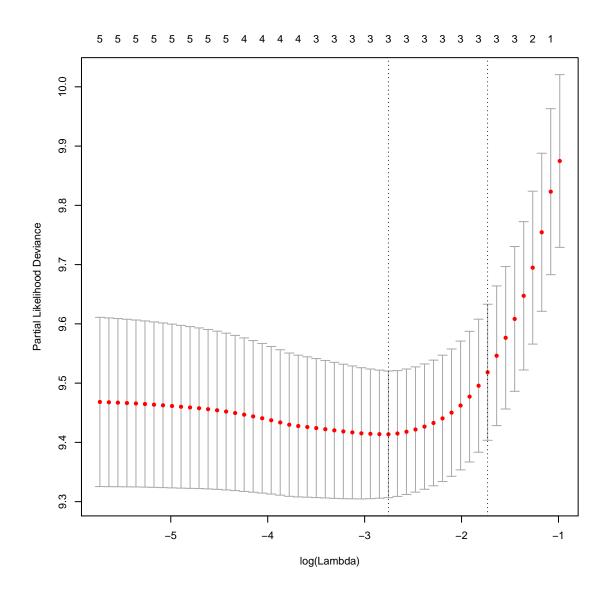
## mg.2 5.568

## mg.3 .

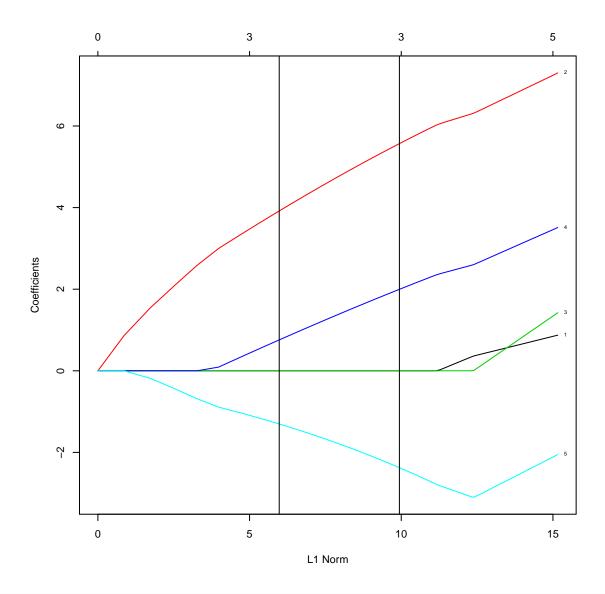
## mg.4 1.997
```

```
## mg.5 -2.374
```

```
plot(glmnet.fit.cv)
```

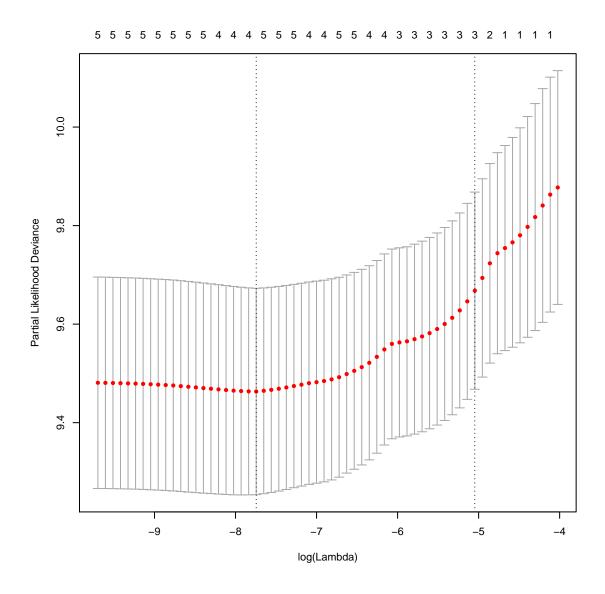


```
plot(glmnet.fit.cv$glmnet.fit, label = TRUE)
abline(v = sum(abs(glmnet.coef.1se)))
abline(v = sum(abs(glmnet.coef.min)))
```

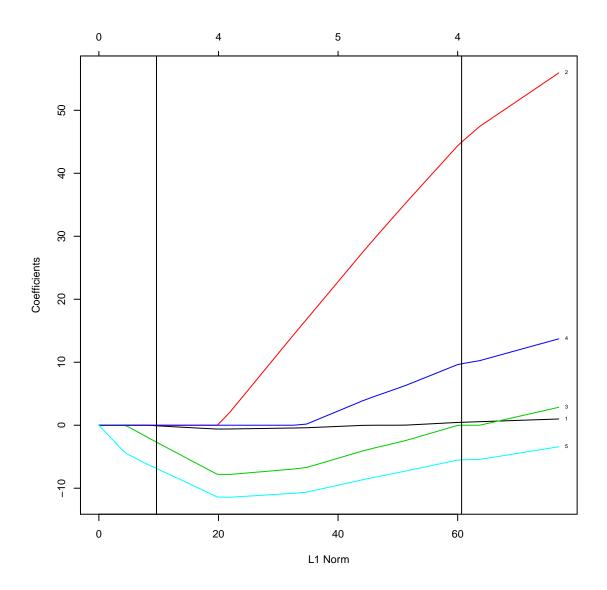


```
## mg.5 -8.7940
```

```
plot(adaglmnet.fit.cv)
```



```
plot(adaglmnet.fit.cv$glmnet.fit, label = TRUE)
abline(v = sum(abs(adaglmnet.coef.1se)))
abline(v = sum(abs(adaglmnet.coef.min)))
```



5 Session information

```
session_info
## R version 3.1.1 (2014-07-10)
## Platform: x86_64-unknown-linux-gnu (64-bit)
##
## locale:
   [1] LC_CTYPE=en_US.iso885915
                                          LC_NUMERIC=C
##
   [3] LC_TIME=en_US.iso885915
                                          LC_COLLATE=en_US.iso885915
##
   [5] LC_MONETARY=en_US.iso885915
                                          LC_MESSAGES=en_US.iso885915
##
   [7] LC_PAPER=en_US.iso885915
                                          LC_NAME=en_US.iso885915
   [9] LC_ADDRESS=en_US.iso885915
                                          LC_TELEPHONE=en_US.iso885915
                                          LC_IDENTIFICATION=en_US.iso885915
## [11] LC_MEASUREMENT=en_US.iso885915
##
## attached base packages:
```

```
## [1] splines
                 parallel methods
                                     stats
                                               graphics grDevices utils
## [8] datasets
                base
## other attached packages:
                                                foreach 1.4.2
## [1] doParallel_1.0.8
                           iterators 1.0.7
## [4] ahaz_1.14
                            survival_2.37-7
                                                stargazer_5.1
## [7] gplots_2.14.2
                            NMF_0.20.4
                                                synchronicity_1.1.4
## [10] bigmemory_4.4.6
                            BH_1.54.0-5
                                                bigmemory.sri_0.1.3
## [13] Biobase_2.26.0
                            BiocGenerics_0.12.1 cluster_1.15.3
## [16] rngtools_1.2.4
                            pkgmaker_0.22
                                                registry_0.2
## [19] RColorBrewer_1.0-5 energy_1.6.2
                                                glmnet_1.9-8
## [22] Matrix_1.1-4
                            glmulti_1.0.7
                                                rJava_0.9-6
## [25] xtable_1.7-4
##
## loaded via a namespace (and not attached):
## [1] bitops_1.0-6
                           boot_1.3-13
                                              caTools_1.17.1
## [4] codetools_0.2-9
                           colorspace_1.2-4
                                              compiler_3.1.1
## [7] digest_0.6.4
                           gdata_2.13.3
                                              ggplot2_1.0.0
## [10] grid_3.1.1
                           gridBase_0.4-7
                                              gtable_0.1.2
                           KernSmooth_2.23-13 lattice_0.20-29
## [13] gtools_3.4.1
## [16] MASS_7.3-35
                           munsell_0.4.2
                                              plyr_1.8.1
## [19] proto_0.3-10
                           Rcpp_0.11.3
                                              reshape2_1.4
## [22] scales_0.2.4
                           stringr_0.6.2
                                              tools_3.1.1
sessionInfo()
## R version 3.1.1 (2014-07-10)
## Platform: x86_64-unknown-linux-gnu (64-bit)
##
## locale:
## [1] LC_CTYPE=en_US.iso885915
                                          LC_NUMERIC=C
## [3] LC_TIME=en_US.iso885915
                                          LC_COLLATE=en_US.iso885915
## [5] LC_MONETARY=en_US.iso885915
                                          LC_MESSAGES=en_US.iso885915
## [7] LC_PAPER=en_US.iso885915
                                          LC_NAME=en_US.iso885915
## [9] LC_ADDRESS=en_US.iso885915
                                          LC_TELEPHONE=en_US.iso885915
## [11] LC_MEASUREMENT=en_US.iso885915
                                         LC_IDENTIFICATION=en_US.iso885915
## attached base packages:
## [1] parallel methods
                          splines
                                               graphics grDevices utils
                                     stats
## [8] datasets base
## other attached packages:
## [1] stargazer_5.1
                            xtable_1.7-4
                                                gplots_2.14.2
## [4] RColorBrewer_1.0-5 glmnet_1.9-8
                                                Matrix_1.1-4
## [7] glmulti_1.0.7
                            rJava_0.9-6
                                                NMF_0.20.4
## [10] synchronicity_1.1.4 bigmemory_4.4.6
                                                BH_1.54.0-5
## [13] bigmemory.sri_0.1.3 Biobase_2.26.0
                                                BiocGenerics_0.12.1
## [16] cluster_1.15.3
                            rngtools_1.2.4
                                                pkgmaker_0.22
## [19] registry_0.2
                            energy_1.6.2
                                                survival_2.37-7
## [22] knitr_1.8
##
## loaded via a namespace (and not attached):
## [1] bitops_1.0-6
                           boot_1.3-13
                                              caTools_1.17.1
## [4] codetools_0.2-9
                        colorspace_1.2-4 digest_0.6.4
```

```
## [7] doParallel_1.0.8
                          evaluate_0.5.5
                                             foreach_1.4.2
## [10] formatR_1.0
                          gdata_2.13.3
                                             ggplot2_1.0.0
## [13] grid_3.1.1
                          gridBase_0.4-7
                                             gtable_0.1.2
## [16] gtools_3.4.1
                          highr_0.4
                                             iterators_1.0.7
## [19] KernSmooth_2.23-13 labeling_0.3
                                             lattice_0.20-29
## [22] MASS_7.3-35
                          munsell_0.4.2
                                             plyr_1.8.1
## [25] proto_0.3-10
                          Rcpp_0.11.3
                                             reshape2_1.4
## [28] scales_0.2.4
                          stringr_0.6.2
                                             tools_3.1.1
```