Homework 4

Charles Hwang

Professor Matthews

STAT 388-001

26 October 2019

Problem 3b

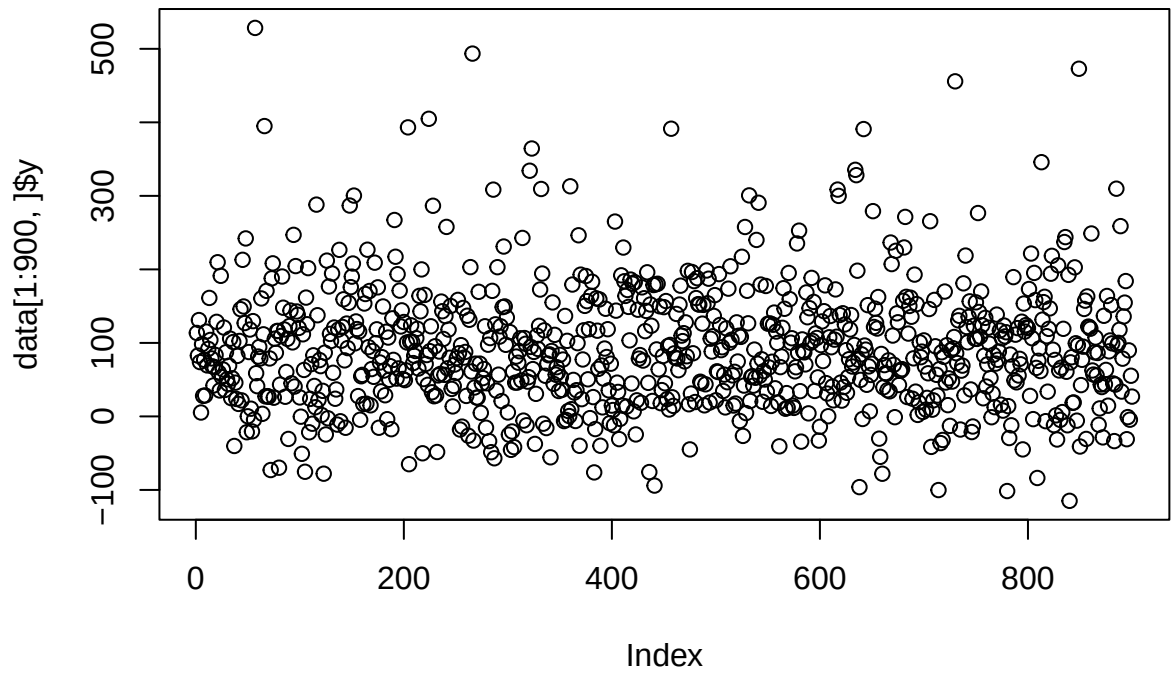
Summary

```
rm(list=ls())
data <- read.csv(file="/Users/newuser/Desktop/Notes/Undergraduate/STAT 338 - Predictive Analytics/Predi
library(corrplot)
library(mgcv)
summary(data[1:900,])
```

```
##        x1                x2                x3                x4
##  Min.   :0.001137   Min.   :0.005158   Min.   :-9.9915   Min.   :0.0008856
##  1st Qu.:0.268251   1st Qu.:2.706508   1st Qu.:-5.2944   1st Qu.:0.2421445
##  Median :0.531032   Median :5.111614   Median :-0.1787   Median :0.5043533
##  Mean   :0.518031   Mean   :5.144425   Mean   :-0.1658   Mean   :0.4914087
##  3rd Qu.:0.762125   3rd Qu.:7.706497   3rd Qu.: 4.9789   3rd Qu.:0.7376653
##  Max.   :0.996996   Max.   :9.995669   Max.   : 9.9919   Max.   :0.9996596
##        x5                x6              x7                  x8
##  Min.   :0.004199   Min.   :4.008   Min.   :0.0007152   Min.   : 0.03929
##  1st Qu.:1.290763   1st Qu.:5.602   1st Qu.:0.2505563   1st Qu.:26.47365
##  Median :2.514286   Median :7.029   Median :0.4800897   Median :51.31043
##  Mean   :2.476138   Mean   :7.017   Mean   :0.4939892   Mean   :51.26421
##  3rd Qu.:3.642716   3rd Qu.:8.479   3rd Qu.:0.7423080   3rd Qu.:76.53738
##  Max.   :4.995992   Max.   :9.983   Max.   :0.9975668   Max.   :99.94112
##        x9                x10               x11               x12
##  Min.   :0.000207   Min.   :-9.98699   Min.   :-2.95948   Min.   :-6.367284
##  1st Qu.:0.958349   1st Qu.:-6.73080   1st Qu.:-0.69950   1st Qu.:-1.381296
##  Median :1.938711   Median :-3.03864   Median :-0.03948   Median :-0.038926
##  Mean   :1.949677   Mean   :-3.37113   Mean   :-0.04704   Mean   : 0.009159
##  3rd Qu.:2.905676   3rd Qu.:-0.07997   3rd Qu.: 0.65961   3rd Qu.: 1.438971
##  Max.   :3.999867   Max.   : 2.99098   Max.   : 2.60437   Max.   : 6.116449
##       x13                x14               x15               x16
##  Min.   :-9.273040   Min.   :-12.72128   Min.   :-30.3208   Min.   :-14.6088
##  1st Qu.:-1.949556   1st Qu.: -2.60289   1st Qu.: -6.4834   1st Qu.: -2.1359
##  Median :-0.003584   Median : -0.04404   Median :  0.2114   Median :  0.5252
##  Mean   :-0.029320   Mean   : -0.05360   Mean   :  0.4482   Mean   :  0.5538
##  3rd Qu.: 1.891750   3rd Qu.:  2.43221   3rd Qu.:  6.6231   3rd Qu.:  3.1833
##  Max.   :10.068275   Max.   : 11.98601   Max.   : 29.1733   Max.   : 15.7934
##       x17               x18               x19               x20
##  Min.   :-6.045   Min.   :-14.44809   Min.   :-2.3776   Min.   :-68.024
```

```
##  1st Qu.: 2.174   1st Qu.: -5.32857   1st Qu.:-0.2386   1st Qu.: -9.836
##  Median : 5.005   Median :  0.08948   Median : 0.4872   Median :  2.475
##  Mean   : 4.958   Mean   : -0.09580   Mean   : 0.4958   Mean   :  2.401
##  3rd Qu.: 7.675   3rd Qu.:  5.06707   3rd Qu.: 1.2007   3rd Qu.: 15.498
##  Max.   :15.500   Max.   : 15.43649   Max.   : 3.6671   Max.   : 65.903
##       x21               x22               x23               x24
##  Min.   :-4.2193   Min.   :-6.9141   Min.   :-72.415   Min.   :-66.988
##  1st Qu.:-0.5168   1st Qu.: 0.6206   1st Qu.:-11.537   1st Qu.:-11.445
##  Median : 0.4168   Median : 2.4765   Median :  2.867   Median :  1.578
##  Mean   : 0.4483   Mean   : 2.4828   Mean   :  2.828   Mean   :  2.205
##  3rd Qu.: 1.3580   3rd Qu.: 4.4398   3rd Qu.: 17.104   3rd Qu.: 16.791
##  Max.   : 4.9335   Max.   :10.2026   Max.   : 86.904   Max.   : 67.439
##       x25                y                 id
##  Min.   :-29.390   Min.   :-114.82   Min.   :  1.0
##  1st Qu.: -5.042   1st Qu.:  35.63   1st Qu.:225.8
##  Median :  1.365   Median :  83.64   Median :450.5
##  Mean   :  1.318   Mean   :  91.43   Mean   :450.5
##  3rd Qu.:  7.753   3rd Qu.: 136.10   3rd Qu.:675.2
##  Max.   : 32.019   Max.   : 528.44   Max.   :900.0
```

```
plot(data[1:900,]$y)
```




```
summary(data[901:1000,])
```

```
##        x1                x2                x3                x4
##  Min.   :0.01266   Min.   :0.2364   Min.   :-9.9702   Min.   :0.007543
##  1st Qu.:0.27790   1st Qu.:2.8799   1st Qu.:-5.8435   1st Qu.:0.226974
##  Median :0.45847   Median :4.5710   Median : 0.2840   Median :0.453825
##  Mean   :0.47820   Mean   :5.0587   Mean   :-0.1958   Mean   :0.480919
##  3rd Qu.:0.66641   3rd Qu.:7.4326   3rd Qu.: 4.6071   3rd Qu.:0.740892
##  Max.   :0.94121   Max.   :9.8401   Max.   : 9.9771   Max.   :0.996088
##
##        x5                x6               x7                x8
##  Min.   :0.02477   Min.   :4.031   Min.   :0.01423   Min.   : 2.074
```

```
##  1st Qu.:1.26559   1st Qu.:5.266   1st Qu.:0.26286   1st Qu.:28.495
##  Median :2.00873   Median :6.682   Median :0.48319   Median :55.725
##  Mean   :2.32876   Mean   :6.834   Mean   :0.49342   Mean   :52.834
##  3rd Qu.:3.52376   3rd Qu.:8.539   3rd Qu.:0.71990   3rd Qu.:80.340
##  Max.   :4.99362   Max.   :9.941   Max.   :0.99830   Max.   :98.325
##
##        x9                x10               x11                x12
##  Min.   :0.002771   Min.   :-9.9287   Min.   :-3.04862   Min.   :-7.003231
##  1st Qu.:1.227045   1st Qu.:-6.6031   1st Qu.:-0.49704   1st Qu.:-1.207122
##  Median :1.952982   Median :-3.6166   Median : 0.05653   Median :-0.054459
##  Mean   :2.014662   Mean   :-3.6224   Mean   : 0.06175   Mean   : 0.002125
##  3rd Qu.:2.821488   3rd Qu.:-0.4678   3rd Qu.: 0.69801   3rd Qu.: 1.291554
##  Max.   :3.961343   Max.   : 2.9236   Max.   : 2.75808   Max.   : 5.055231
##
##       x13                x14               x15                x16
##  Min.   :-8.01982   Min.   :-9.7323   Min.   :-21.6103   Min.   :-11.2229
##  1st Qu.:-2.29355   1st Qu.:-3.1131   1st Qu.: -7.5668   1st Qu.: -1.9800
##  Median :-0.12163   Median : 0.1209   Median :  0.1105   Median :  0.8090
##  Mean   : 0.07993   Mean   :-0.0095   Mean   :  0.2030   Mean   :  0.5109
##  3rd Qu.: 2.55936   3rd Qu.: 2.8922   3rd Qu.:  7.2836   3rd Qu.:  2.9089
##  Max.   : 8.93216   Max.   :11.0597   Max.   : 22.2581   Max.   :  9.9715
##
##       x17              x18                x19                x20
##  Min.   :-1.521   Min.   :-12.5793   Min.   :-2.22969   Min.   :-64.920
##  1st Qu.: 3.129   1st Qu.: -5.4963   1st Qu.:-0.03434   1st Qu.: -6.915
##  Median : 5.239   Median : -0.6325   Median : 0.65349   Median :  5.397
##  Mean   : 5.582   Mean   : -0.4093   Mean   : 0.60998   Mean   :  3.068
##  3rd Qu.: 7.854   3rd Qu.:  4.5971   3rd Qu.: 1.27821   3rd Qu.: 17.296
##  Max.   :15.620   Max.   : 12.6575   Max.   : 3.42094   Max.   : 41.390
##
##       x21               x22               x23               x24
##  Min.   :-3.9413   Min.   :-7.957   Min.   :-72.880   Min.   :-69.257
##  1st Qu.:-0.3444   1st Qu.: 0.408   1st Qu.:-12.325   1st Qu.: -5.723
##  Median : 0.4707   Median : 2.397   Median :  7.824   Median :  9.423
##  Mean   : 0.5411   Mean   : 2.252   Mean   :  3.352   Mean   :  3.212
##  3rd Qu.: 1.6650   3rd Qu.: 3.775   3rd Qu.: 19.274   3rd Qu.: 18.333
##  Max.   : 5.0651   Max.   :10.165   Max.   : 45.417   Max.   : 54.109
##
##       x25                y              id
##  Min.   :-20.115   Min.   : NA   Min.   : 901.0
##  1st Qu.: -4.079   1st Qu.: NA   1st Qu.: 925.8
##  Median :  1.683   Median : NA   Median : 950.5
##  Mean   :  2.147   Mean   :NaN   Mean   : 950.5
##  3rd Qu.:  7.686   3rd Qu.: NA   3rd Qu.: 975.2
##  Max.   : 30.430   Max.   : NA   Max.   :1000.0
##                    NA's   :100
```

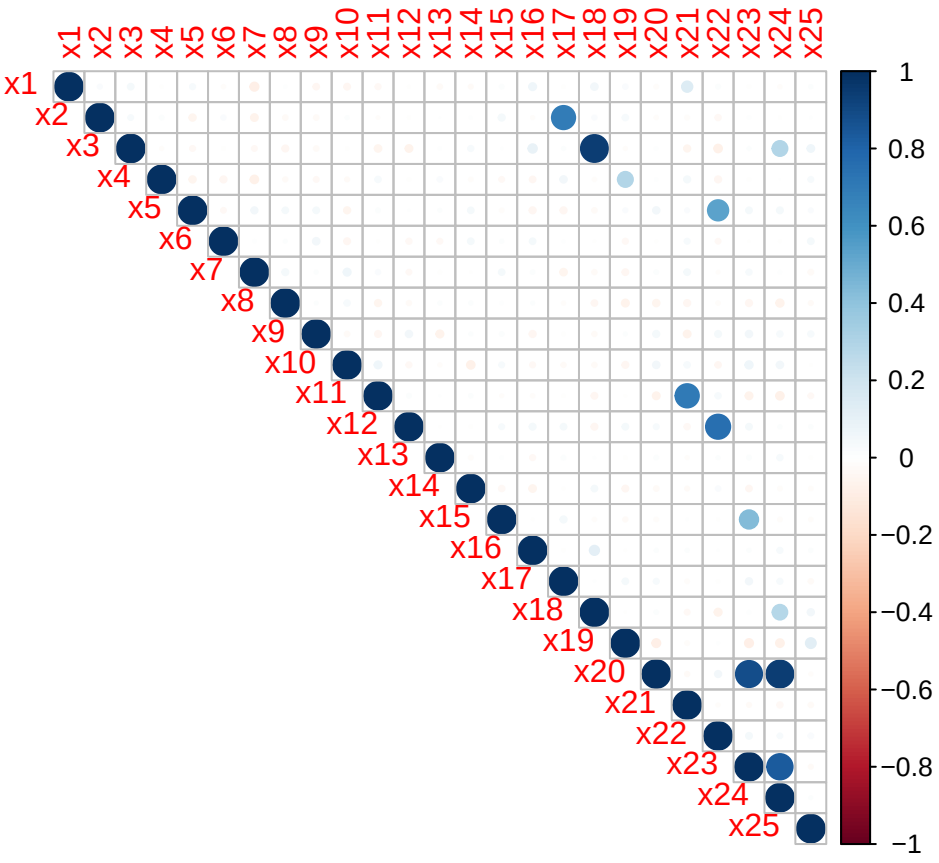
```
round(cor(data[1:900,1:25]),2)
```

```
##        x1    x2    x3    x4    x5    x6    x7    x8    x9   x10   x11   x12
## x1   1.00  0.02  0.04 -0.01  0.03 -0.02 -0.09  0.01 -0.04 -0.04 -0.03 -0.01
## x2   0.02  1.00  0.03  0.02 -0.05  0.02 -0.06 -0.03 -0.03  0.01 -0.01 -0.01
## x3   0.04  0.03  1.00 -0.01 -0.03 -0.01 -0.03 -0.05 -0.03 -0.01 -0.06 -0.06
## x4  -0.01  0.02 -0.01  1.00 -0.05 -0.05 -0.08 -0.02 -0.03 -0.02  0.04 -0.01
## x5   0.03 -0.05 -0.03 -0.05  1.00 -0.03  0.05  0.04  0.04 -0.05  0.02  0.00
```

```
## x6  -0.02  0.02 -0.01 -0.05 -0.03  1.00  0.01  0.01  0.06 -0.05  0.00 -0.04
## x7  -0.09 -0.06 -0.03 -0.08  0.05  0.01  1.00  0.05  0.01  0.07  0.03 -0.03
## x8   0.01 -0.03 -0.05 -0.02  0.04  0.01  0.05  1.00  0.01  0.04 -0.05 -0.02
## x9  -0.04 -0.03 -0.03 -0.03  0.04  0.06  0.01  0.01  1.00 -0.03 -0.05  0.05
## x10 -0.04  0.01 -0.01 -0.02 -0.05 -0.05  0.07  0.04 -0.03  1.00  0.06 -0.03
## x11 -0.03 -0.01 -0.06  0.04  0.02  0.00  0.03 -0.05 -0.05  0.06  1.00  0.02
## x12 -0.01 -0.01 -0.06 -0.01  0.00 -0.04 -0.03 -0.02  0.05 -0.03  0.02  1.00
## x13 -0.03  0.00  0.01  0.03  0.01 -0.03 -0.01  0.01 -0.06 -0.01  0.00  0.00
## x14 -0.02  0.01  0.04 -0.02  0.04  0.04  0.00  0.02  0.01 -0.08  0.02 -0.01
## x15  0.02  0.05  0.00 -0.03 -0.02  0.01  0.04  0.01  0.02  0.03 -0.01  0.03
## x16  0.06 -0.01  0.09 -0.04 -0.05  0.04  0.02  0.01 -0.05 -0.03 -0.02  0.05
## x17 -0.01  0.70  0.01  0.06 -0.05  0.01 -0.06 -0.01 -0.01 -0.02 -0.01  0.04
## x18  0.06  0.02  0.94 -0.02 -0.03  0.00 -0.03 -0.04 -0.03 -0.03 -0.04 -0.05
## x19  0.03  0.04 -0.01  0.29  0.02 -0.03 -0.04 -0.06  0.01 -0.05  0.00  0.04
## x20 -0.02 -0.02  0.01  0.01  0.05  0.00  0.02 -0.06  0.05  0.06 -0.06  0.03
## x21  0.14 -0.02 -0.05  0.05  0.02  0.03  0.04 -0.05 -0.06  0.04  0.71 -0.04
## x22  0.02 -0.04 -0.08 -0.05  0.54 -0.03 -0.01 -0.02  0.05 -0.04  0.04  0.76
## x23 -0.01  0.01  0.01 -0.01  0.04  0.00  0.04 -0.05  0.05  0.07 -0.06  0.04
## x24 -0.01  0.00  0.29  0.00  0.05  0.00  0.01 -0.07  0.04  0.06 -0.07  0.01
## x25  0.01  0.00  0.06  0.03  0.02  0.04 -0.03 -0.02  0.04  0.00 -0.03  0.02
##       x13   x14   x15   x16   x17   x18   x19   x20   x21   x22   x23   x24
## x1  -0.03 -0.02  0.02  0.06 -0.01  0.06  0.03 -0.02  0.14  0.02 -0.01 -0.01
## x2   0.00  0.01  0.05 -0.01  0.70  0.02  0.04 -0.02 -0.02 -0.04  0.01  0.00
## x3   0.01  0.04  0.00  0.09  0.01  0.94 -0.01  0.01 -0.05 -0.08  0.01  0.29
## x4   0.03 -0.02 -0.03 -0.04  0.06 -0.02  0.29  0.01  0.05 -0.05 -0.01  0.00
## x5   0.01  0.04 -0.02 -0.05 -0.05 -0.03  0.02  0.05  0.02  0.54  0.04  0.05
## x6  -0.03  0.04  0.01  0.04  0.01  0.00 -0.03  0.00  0.03 -0.03  0.00  0.00
## x7  -0.01  0.00  0.04  0.02 -0.06 -0.03 -0.04  0.02  0.04 -0.01  0.04  0.01
## x8   0.01  0.02  0.01  0.01 -0.01 -0.04 -0.06 -0.06 -0.05 -0.02 -0.05 -0.07
## x9  -0.06  0.01  0.02 -0.05 -0.01 -0.03  0.01  0.05 -0.06  0.05  0.05  0.04
## x10 -0.01 -0.08  0.03 -0.03 -0.02 -0.03 -0.05  0.06  0.04 -0.04  0.07  0.06
## x11  0.00  0.02 -0.01 -0.02 -0.01 -0.04  0.00 -0.06  0.71  0.04 -0.06 -0.07
## x12  0.00 -0.01  0.03  0.05  0.04 -0.05  0.04  0.03 -0.04  0.76  0.04  0.01
## x13  1.00 -0.02  0.01 -0.03  0.02  0.00 -0.01  0.02 -0.01  0.01  0.02  0.04
## x14 -0.02  1.00 -0.04 -0.06 -0.01  0.04 -0.05 -0.03 -0.02  0.03 -0.04 -0.02
## x15  0.01 -0.04  1.00  0.00  0.05 -0.02 -0.02 -0.02  0.00  0.00  0.44 -0.02
## x16 -0.03 -0.06  0.00  1.00 -0.01  0.11  0.00  0.01  0.01  0.00  0.01  0.03
## x17  0.02 -0.01  0.05 -0.01  1.00  0.00  0.04  0.03 -0.02 -0.01  0.05  0.03
## x18  0.00  0.04 -0.02  0.11  0.00  1.00  0.00  0.02 -0.03 -0.07  0.01  0.28
## x19 -0.01 -0.05 -0.02  0.00  0.04  0.00  1.00 -0.08 -0.02  0.01 -0.09 -0.08
## x20  0.02 -0.03 -0.02  0.01  0.03  0.02 -0.08  1.00 -0.02  0.05  0.89  0.94
## x21 -0.01 -0.02  0.00  0.01 -0.02 -0.03 -0.02 -0.02  1.00 -0.01 -0.02 -0.04
## x22  0.01  0.03  0.00  0.00 -0.01 -0.07  0.01  0.05 -0.01  1.00  0.04  0.03
## x23  0.02 -0.04  0.44  0.01  0.05  0.01 -0.09  0.89 -0.02  0.04  1.00  0.83
## x24  0.04 -0.02 -0.02  0.03  0.03  0.28 -0.08  0.94 -0.04  0.03  0.83  1.00
## x25 -0.01 -0.02 -0.02  0.01 -0.03  0.05  0.13 -0.01 -0.03  0.03 -0.02  0.01
##       x25
## x1   0.01
## x2   0.00
## x3   0.06
## x4   0.03
## x5   0.02
## x6   0.04
## x7  -0.03
```

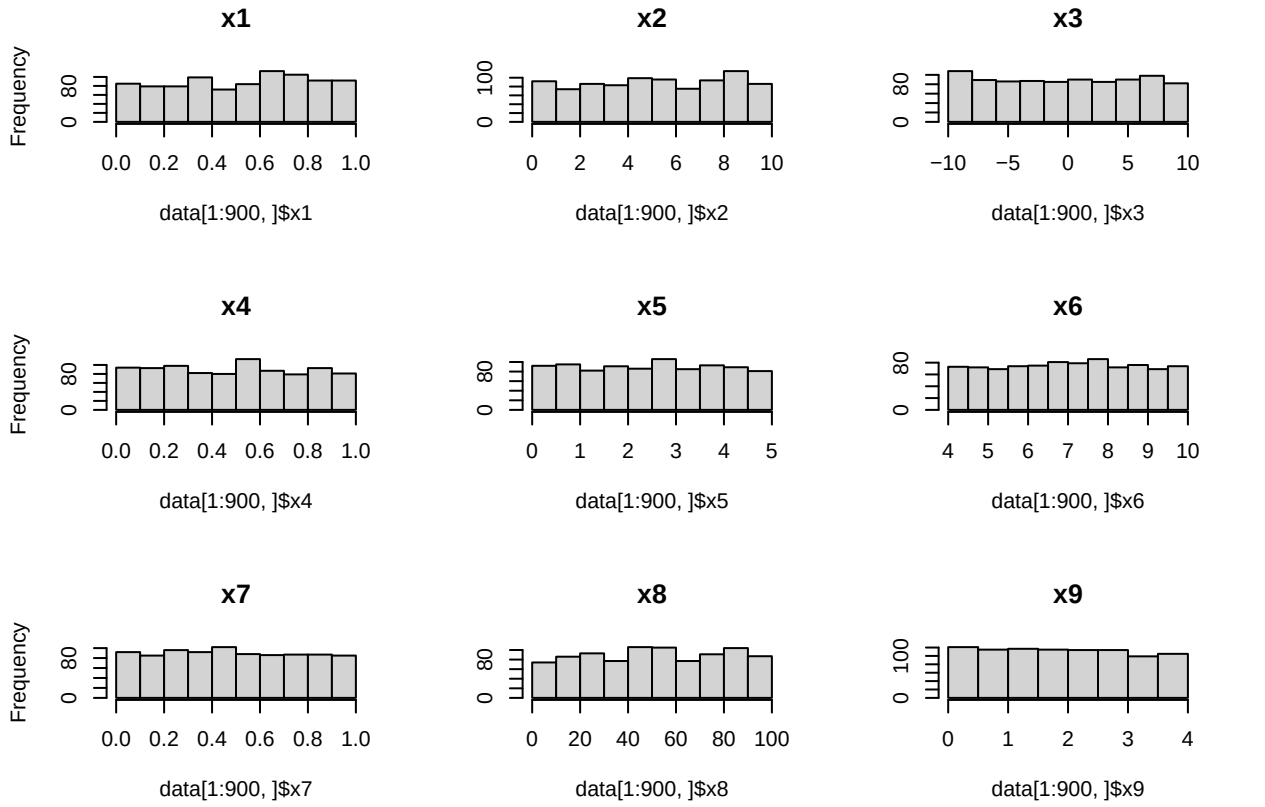
```
## x8  -0.02
## x9   0.04
## x10  0.00
## x11 -0.03
## x12  0.02
## x13 -0.01
## x14 -0.02
## x15 -0.02
## x16  0.01
## x17 -0.03
## x18  0.05
## x19  0.13
## x20 -0.01
## x21 -0.03
## x22  0.03
## x23 -0.02
## x24  0.01
## x25  1.00
```

```
corrplot(cor(data[1:900,1:25]),type="upper")
```

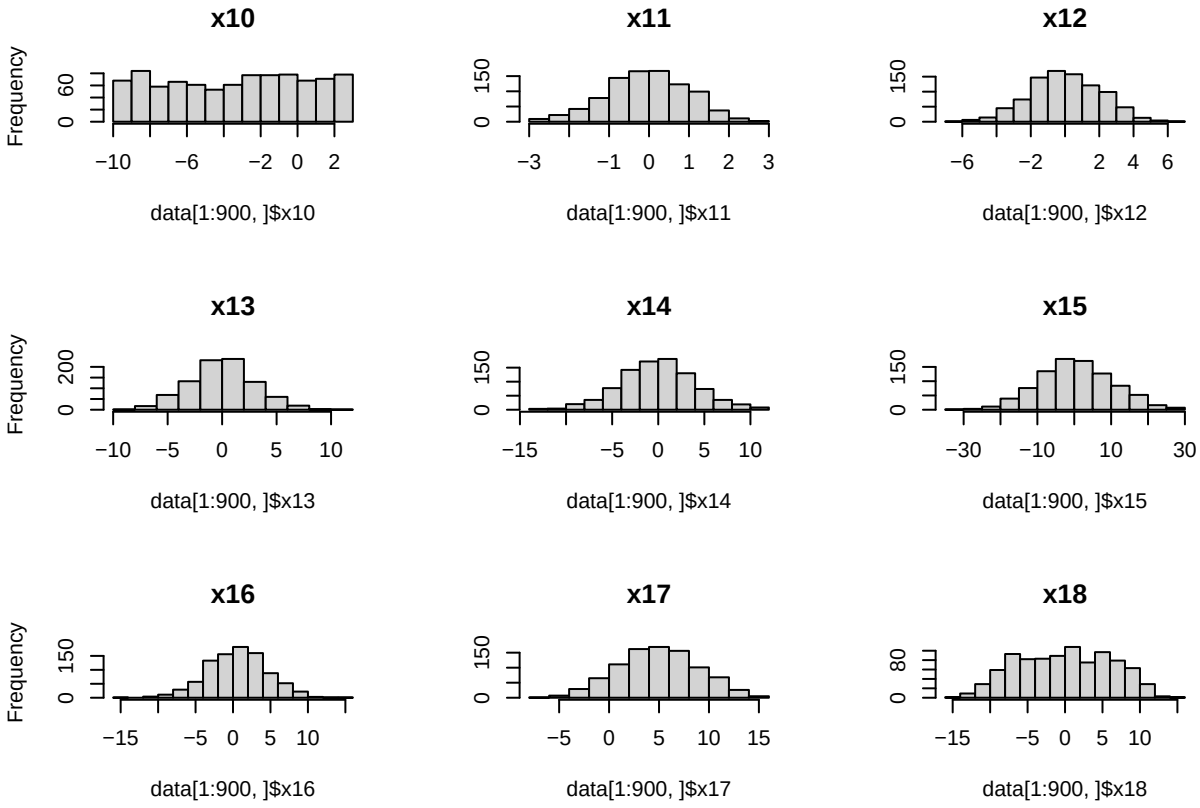



```
par(mfrow=c(3,3))
hist(data[1:900,]$x1,main="x1")
hist(data[1:900,]$x2,main="x2",ylab="")
hist(data[1:900,]$x3,main="x3",ylab="")
hist(data[1:900,]$x4,main="x4")
hist(data[1:900,]$x5,main="x5",ylab="")
```

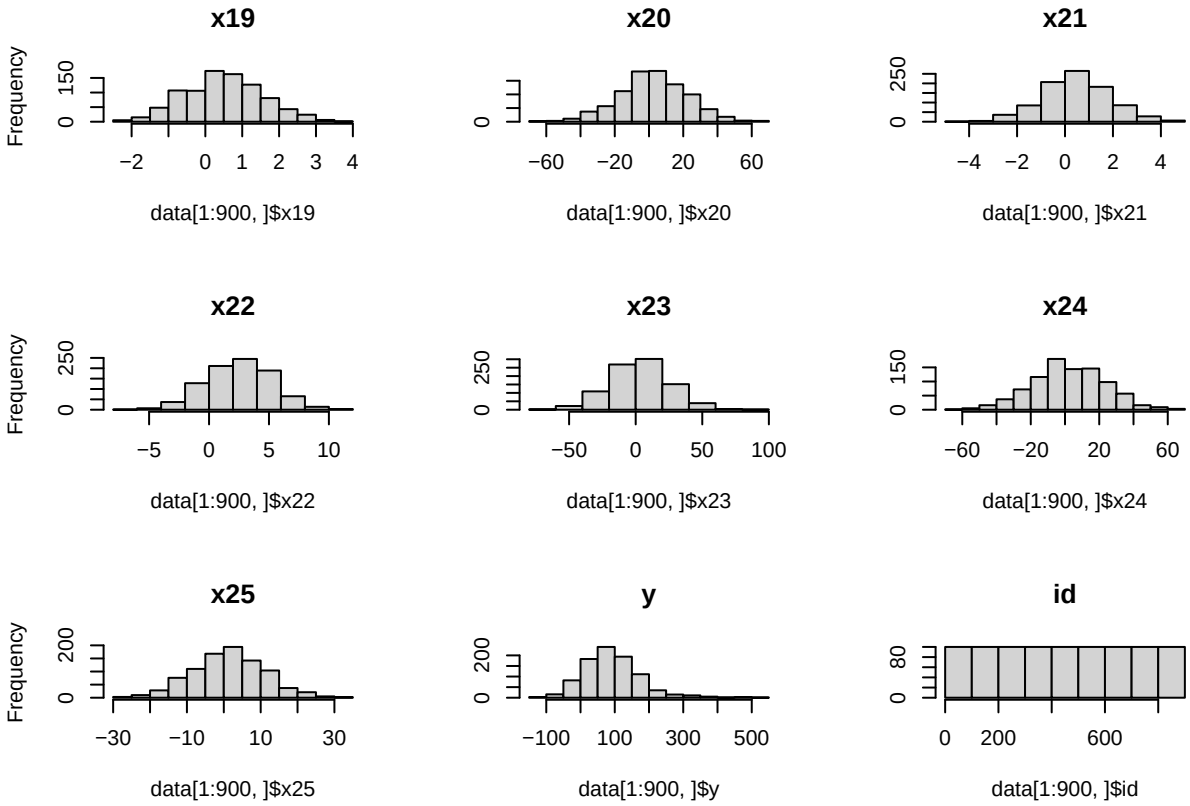
```
hist(data[1:900,]$x6,main="x6",ylab="")
hist(data[1:900,]$x7,main="x7")
hist(data[1:900,]$x8,main="x8",ylab="")
hist(data[1:900,]$x9,main="x9",ylab="")
```




```
par(mfrow=c(3,3))
hist(data[1:900,]$x10,main="x10")
hist(data[1:900,]$x11,main="x11",ylab="")
hist(data[1:900,]$x12,main="x12",ylab="")
hist(data[1:900,]$x13,main="x13")
hist(data[1:900,]$x14,main="x14",ylab="")
hist(data[1:900,]$x15,main="x15",ylab="")
hist(data[1:900,]$x16,main="x16")
hist(data[1:900,]$x17,main="x17",ylab="")
hist(data[1:900,]$x18,main="x18",ylab="")
```




```
par(mfrow=c(3,3))
hist(data[1:900,]$x19,main="x19")
hist(data[1:900,]$x20,main="x20",ylab="")
hist(data[1:900,]$x21,main="x21",ylab="")
hist(data[1:900,]$x22,main="x22")
hist(data[1:900,]$x23,main="x23",ylab="")
hist(data[1:900,]$x24,main="x24",ylab="")
hist(data[1:900,]$x25,main="x25")
hist(data[1:900,]$y,main="y",ylab="")
hist(data[1:900,]$id,main="id",ylab="")
```

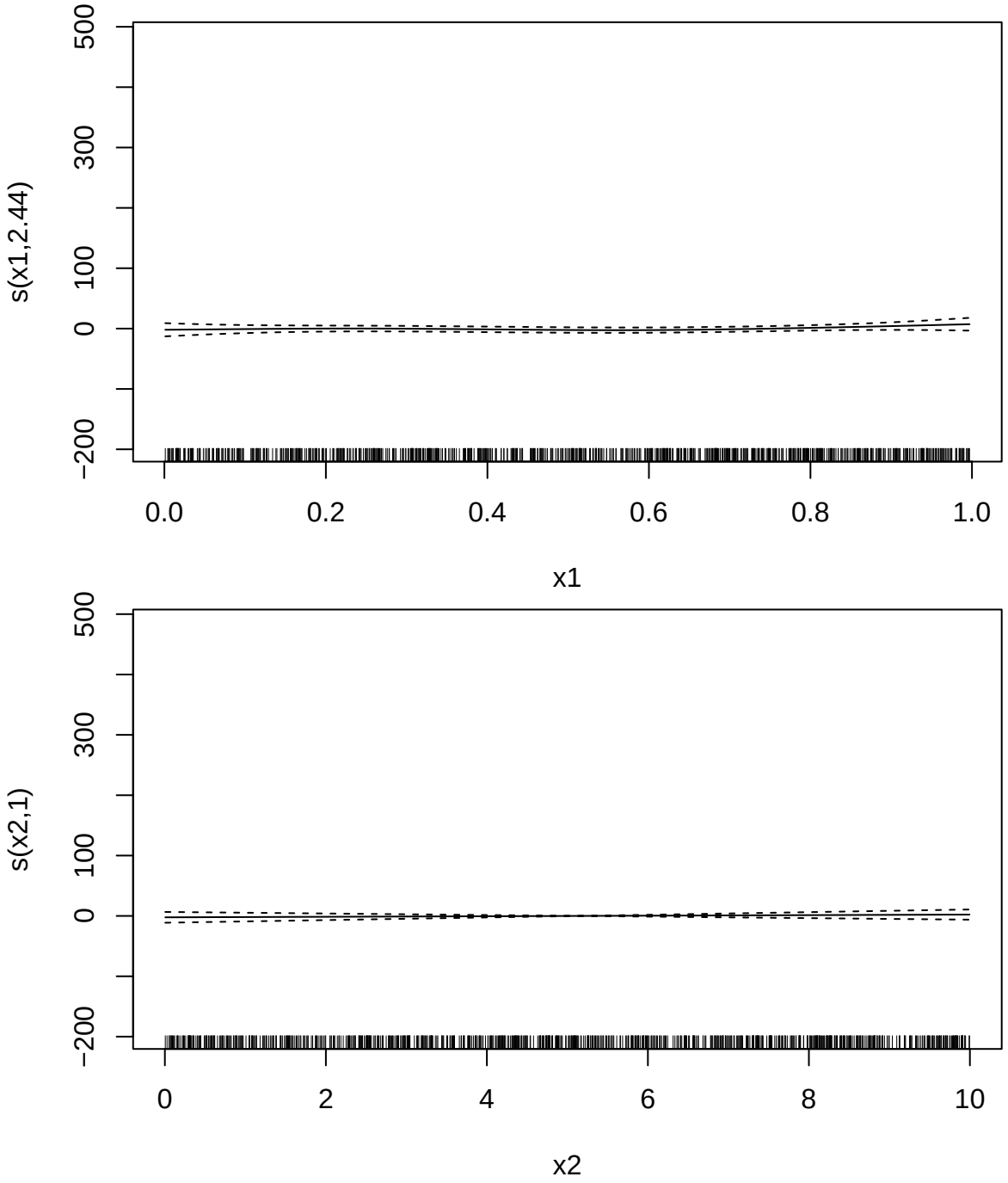


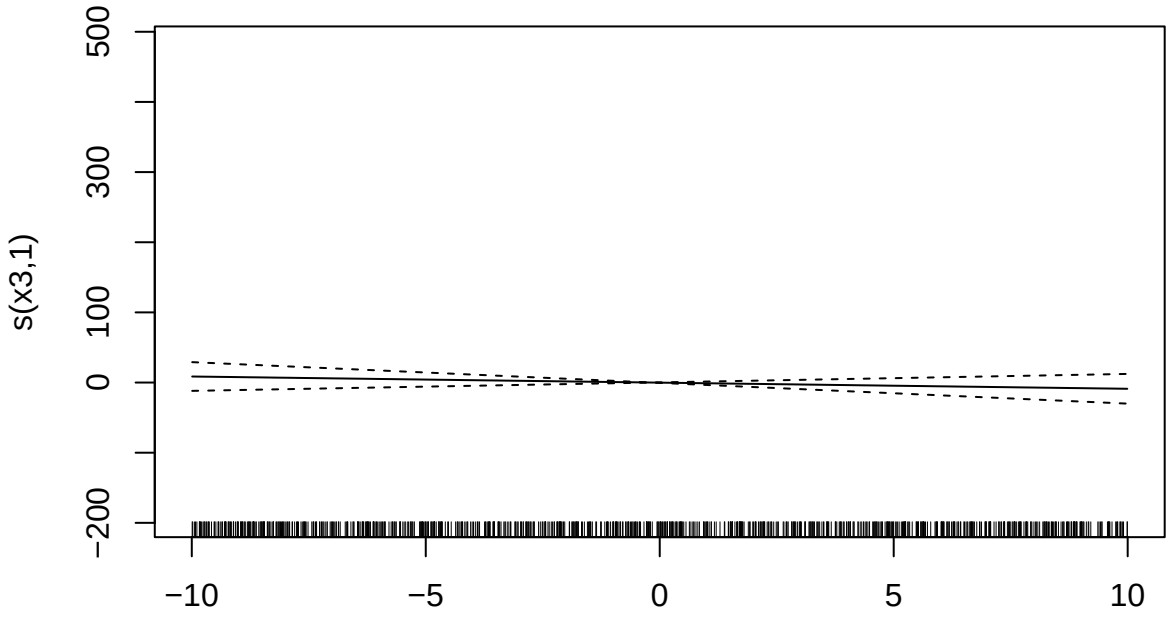

```
summary(gam(y~x1+x2+x3+x4+x5+x6+x7+x8+x9+x10+x11+x12+x13+x14+x15+x16+x17+x18+x19+x20+x21+x22+x23+x24+x25
```

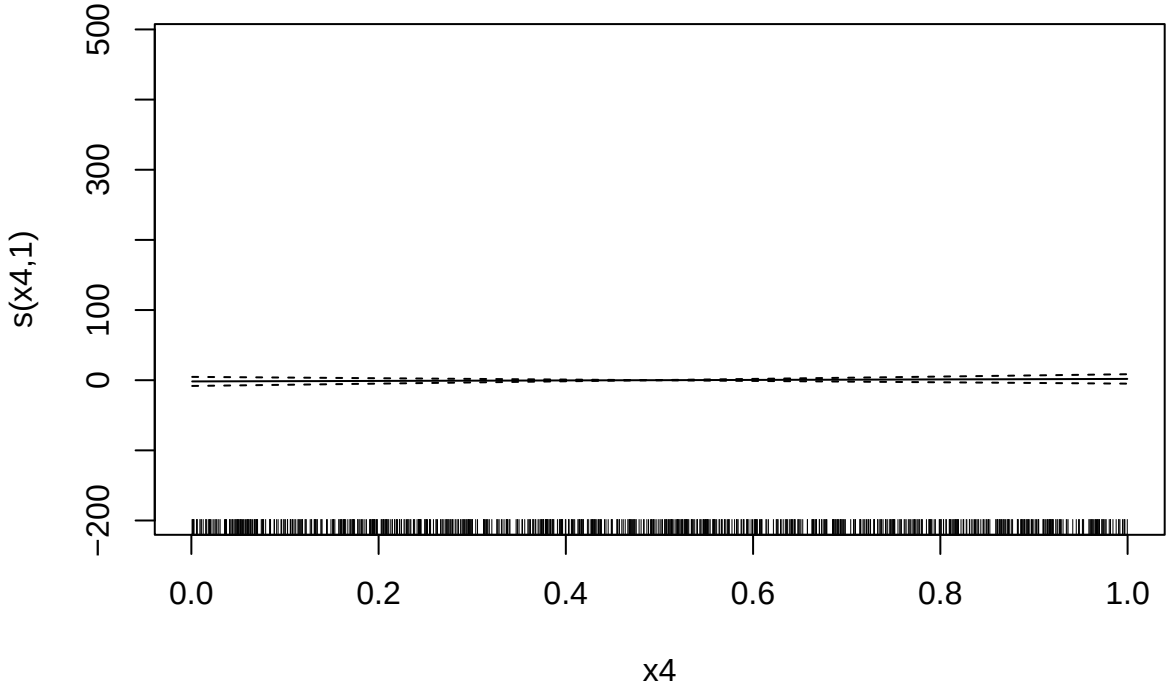
```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 + x10 + x11 +
##     x12 + x13 + x14 + x15 + x16 + x17 + x18 + x19 + x20 + x21 +
##     x22 + x23 + x24 + x25
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 109.195206  17.869880   6.111 1.49e-09 ***
## x1            0.991487  10.055703   0.099   0.9215
## x2            1.035463   1.345916   0.769   0.4419
## x3           -0.225987   1.615230  -0.140   0.8888
## x4            4.759418  10.223977   0.466   0.6417
## x5            1.624840   3.478146   0.467   0.6405
## x6           -1.384660   1.644101  -0.842   0.3999
## x7           -8.975720   9.828709  -0.913   0.3614
## x8           -0.009849   0.099147  -0.099   0.9209
## x9           -3.544468   2.447110  -1.448   0.1479
## x10           1.061090   0.739135   1.436   0.1515
## x11          11.029843   3.995994   2.760   0.0059 **
## x12           0.259308   3.141814   0.083   0.9342
## x13           0.215110   0.946829   0.227   0.8203
## x14           0.353552   0.689648   0.513   0.6083
```

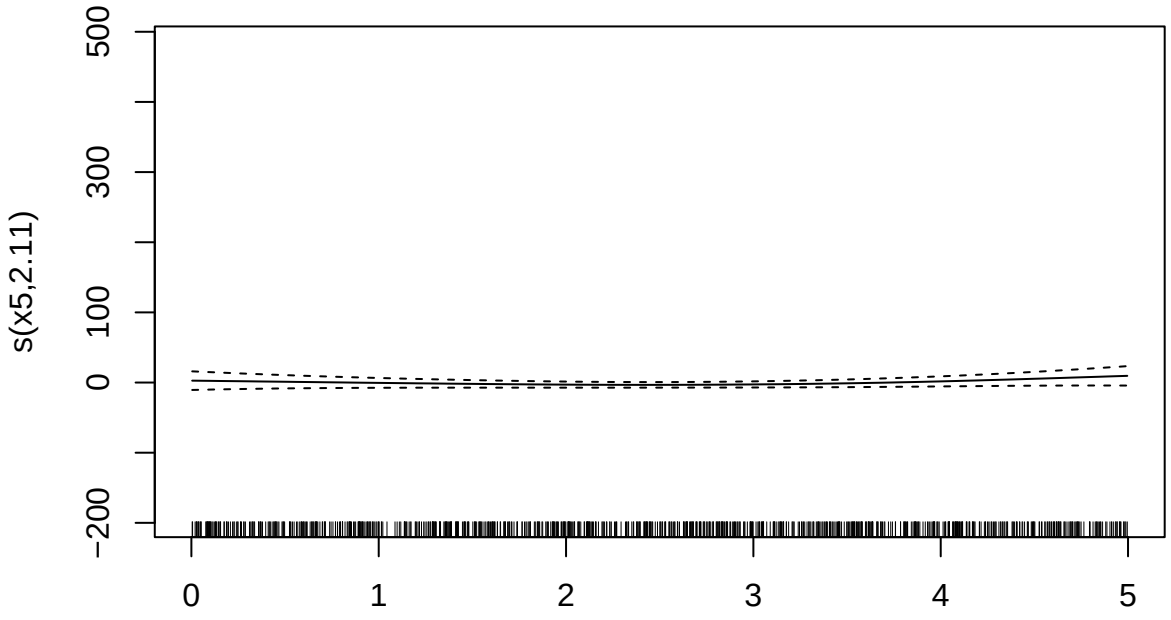
```
## x15           0.364214   1.425284   0.256   0.7984
## x16           0.420726   0.713721   0.589   0.5557
## x17          -0.310758   0.988765  -0.314   0.7534
## x18           0.817770   1.380480   0.592   0.5537
## x19          -0.868506   2.815115  -0.309   0.7578
## x20           0.276433   1.572945   0.176   0.8605
## x21          -2.514781   2.871398  -0.876   0.3814
## x22          -0.853594   2.802647  -0.305   0.7608
## x23          -0.294202   1.387921  -0.212   0.8322
## x24           0.522797   0.710481   0.736   0.4620
## x25          -0.307286   0.290130  -1.059   0.2898
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## R-sq.(adj) =  0.0154   Deviance explained = 4.28%
## GCV = 7076.5  Scale est. = 6872      n = 900
```

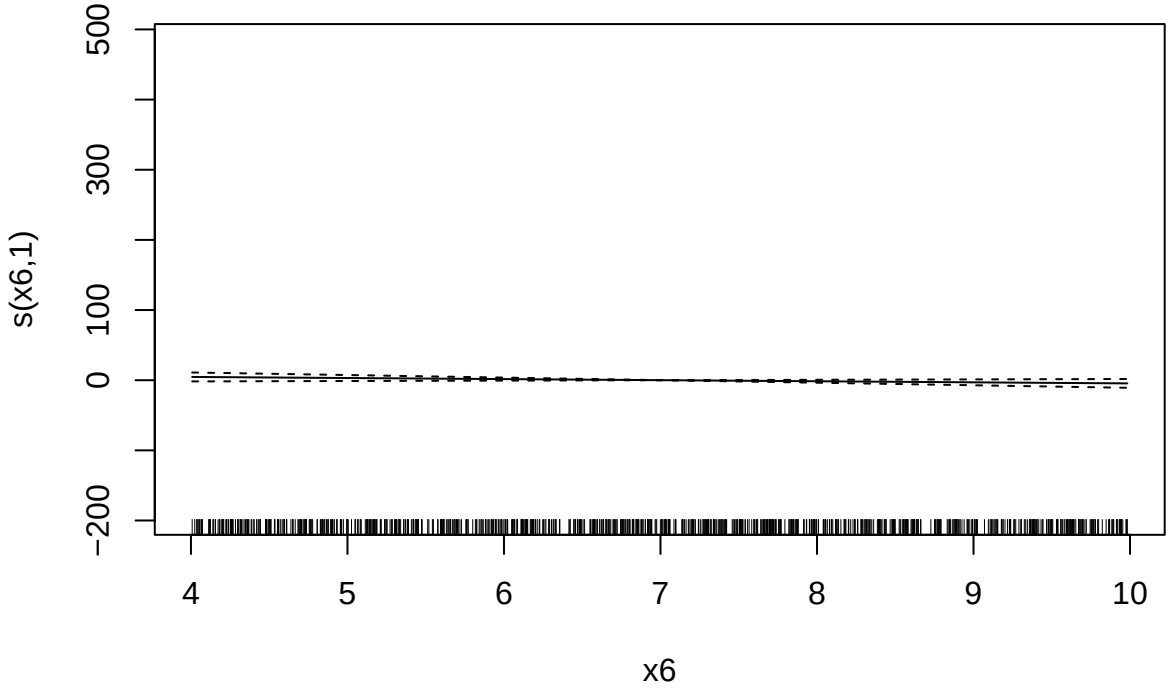
Execution

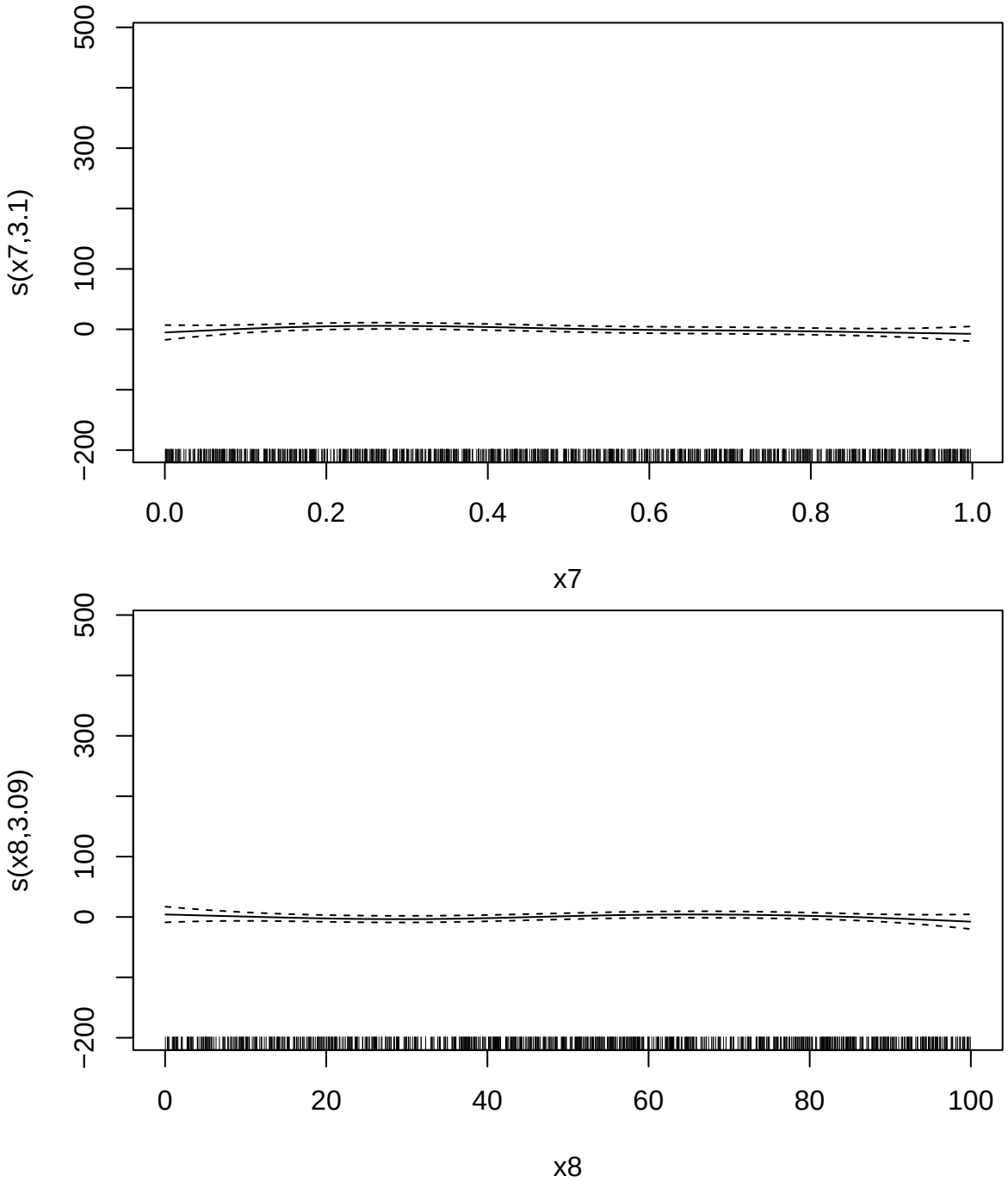

s(x1,2.44)
500
300
100
0
−200
0.0
0.2
0.4
0.6
0.8
1.0
x1
s(x2,1)
500
300
100
0
−200
0
2
4
6
8
10
x2

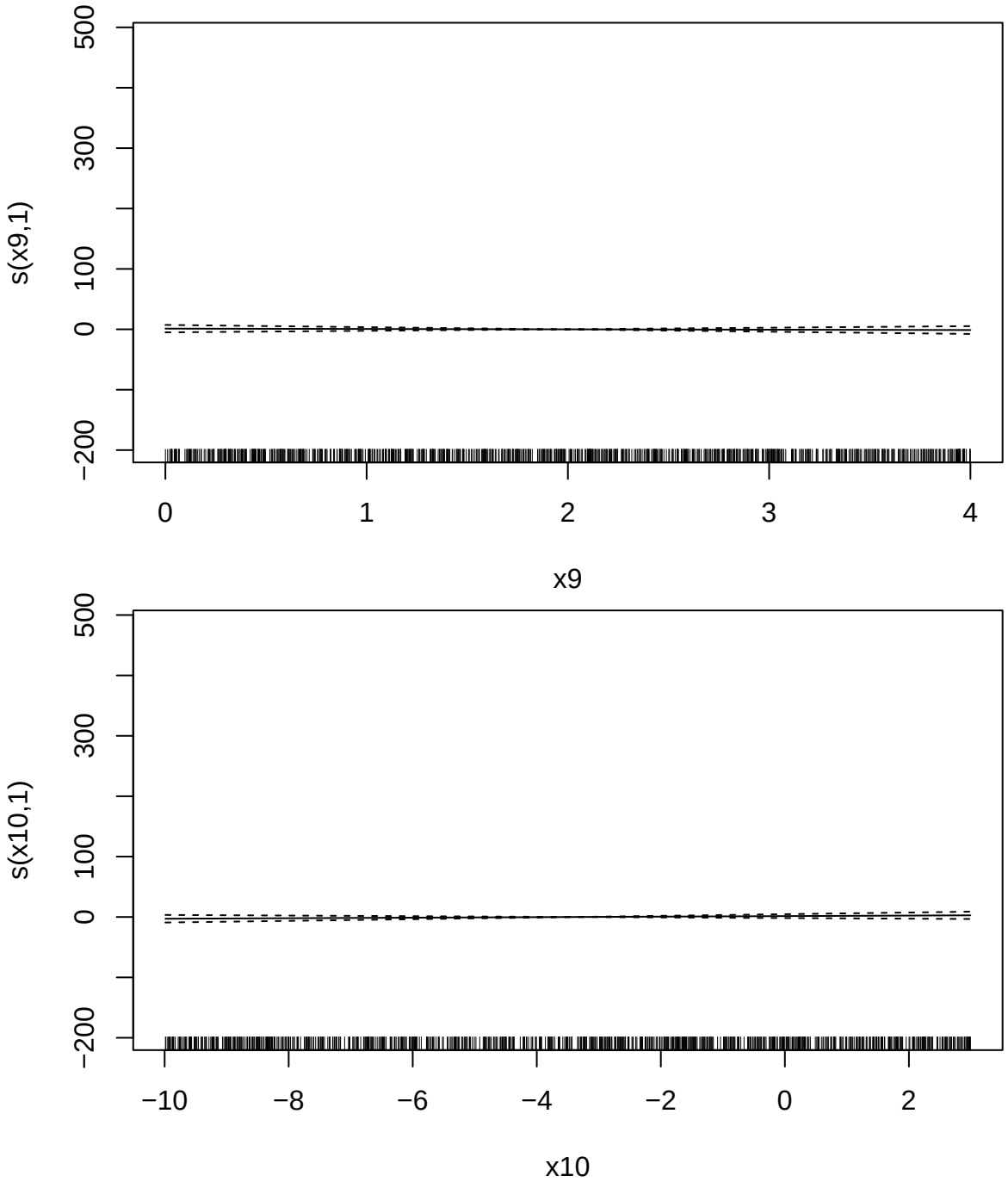

s(x3,1)
500
300
100
0
−200
−10
−5
0
5
10
x3

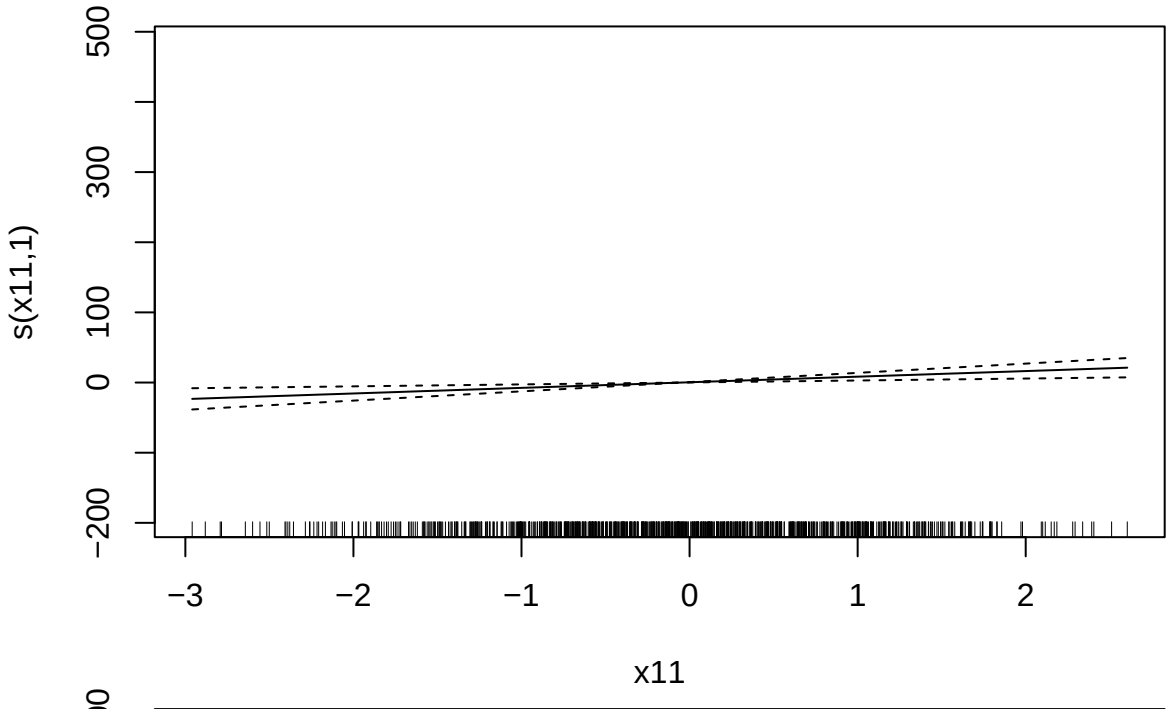

s(x4,1)
500
300
100
0
−200
0.0
0.2
0.4
0.6
0.8
1.0
x4

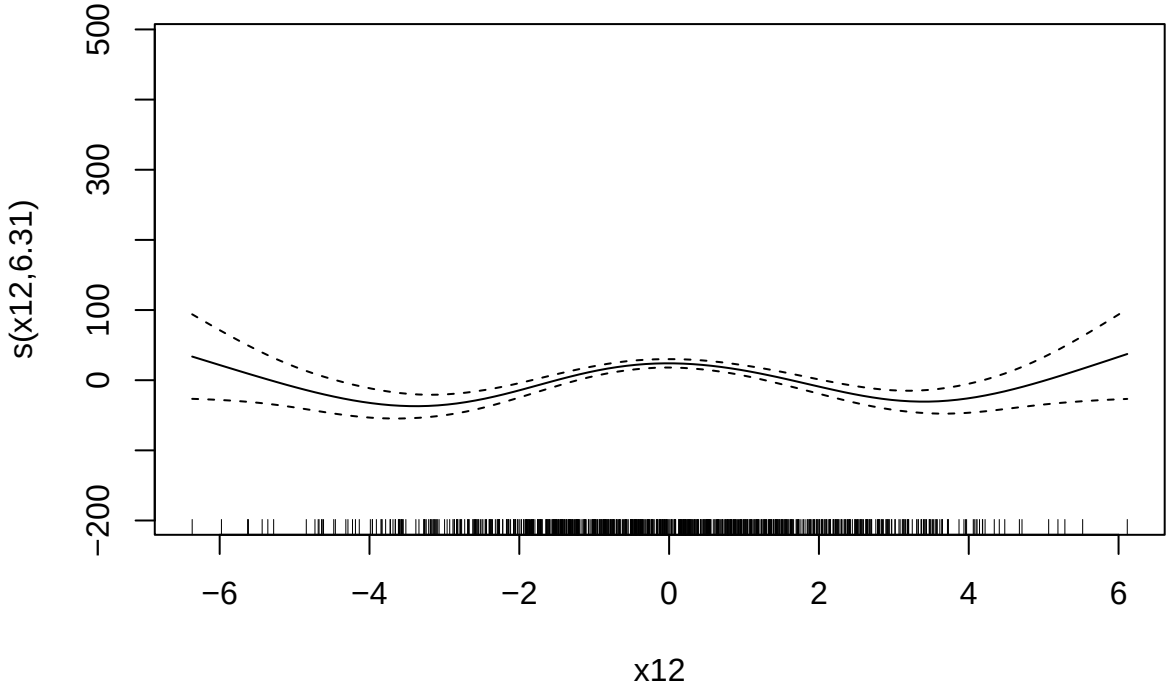

s(x5,2.11)
500
300
100
0
−200
0
1
2
3
4
5
x5

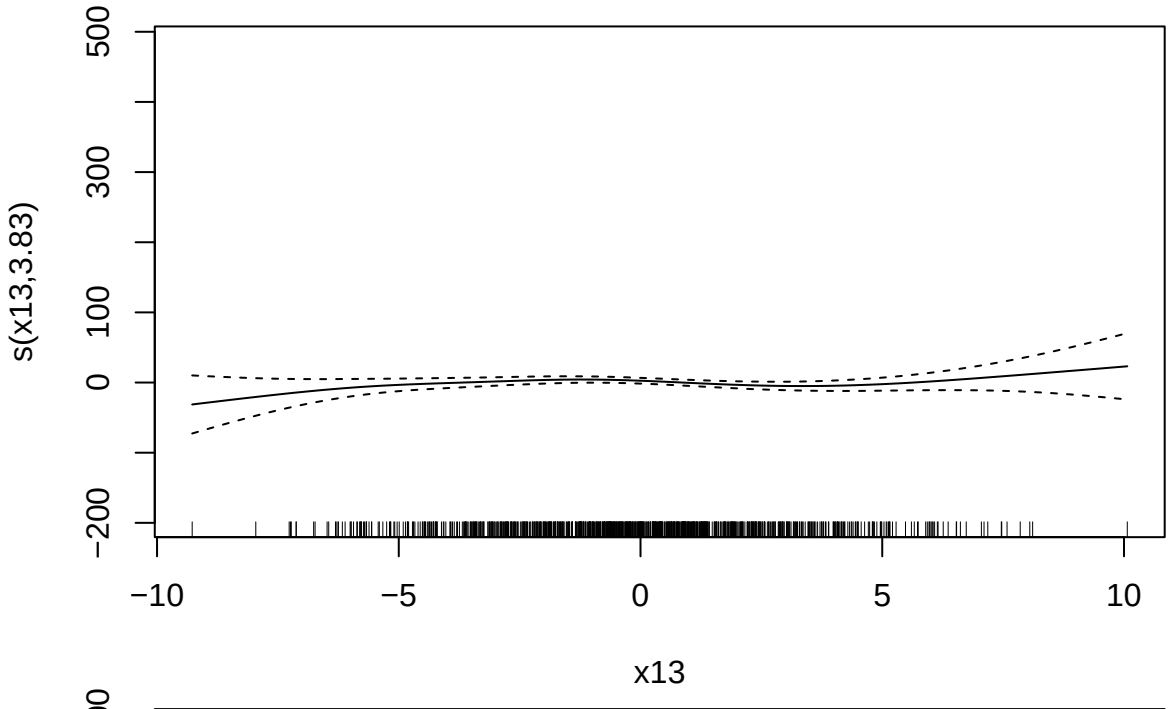

s(x6,1)
500
300
100
0
−200
4
5
6
7
8
9
10
x6

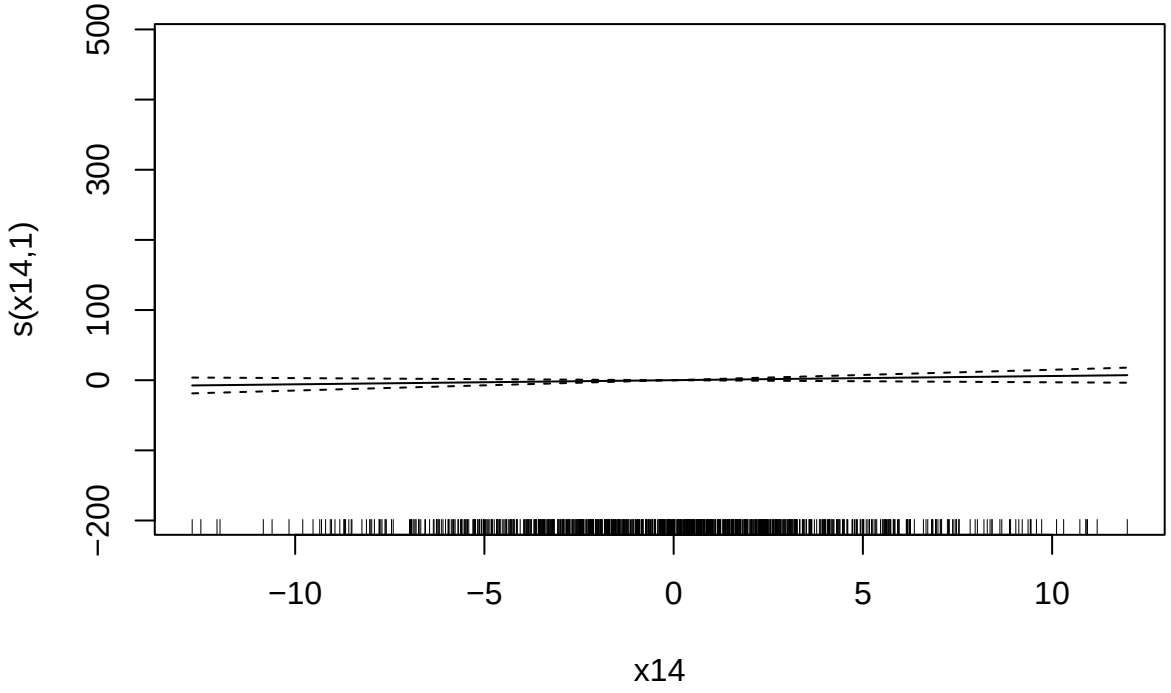

s(x7,3.1)
500
300
100
0
−200
0.0
0.2
0.4
0.6
0.8
1.0
x7
s(x8,3.09)
500
300
100
0
−200
0
20
40
60
80
100
x8

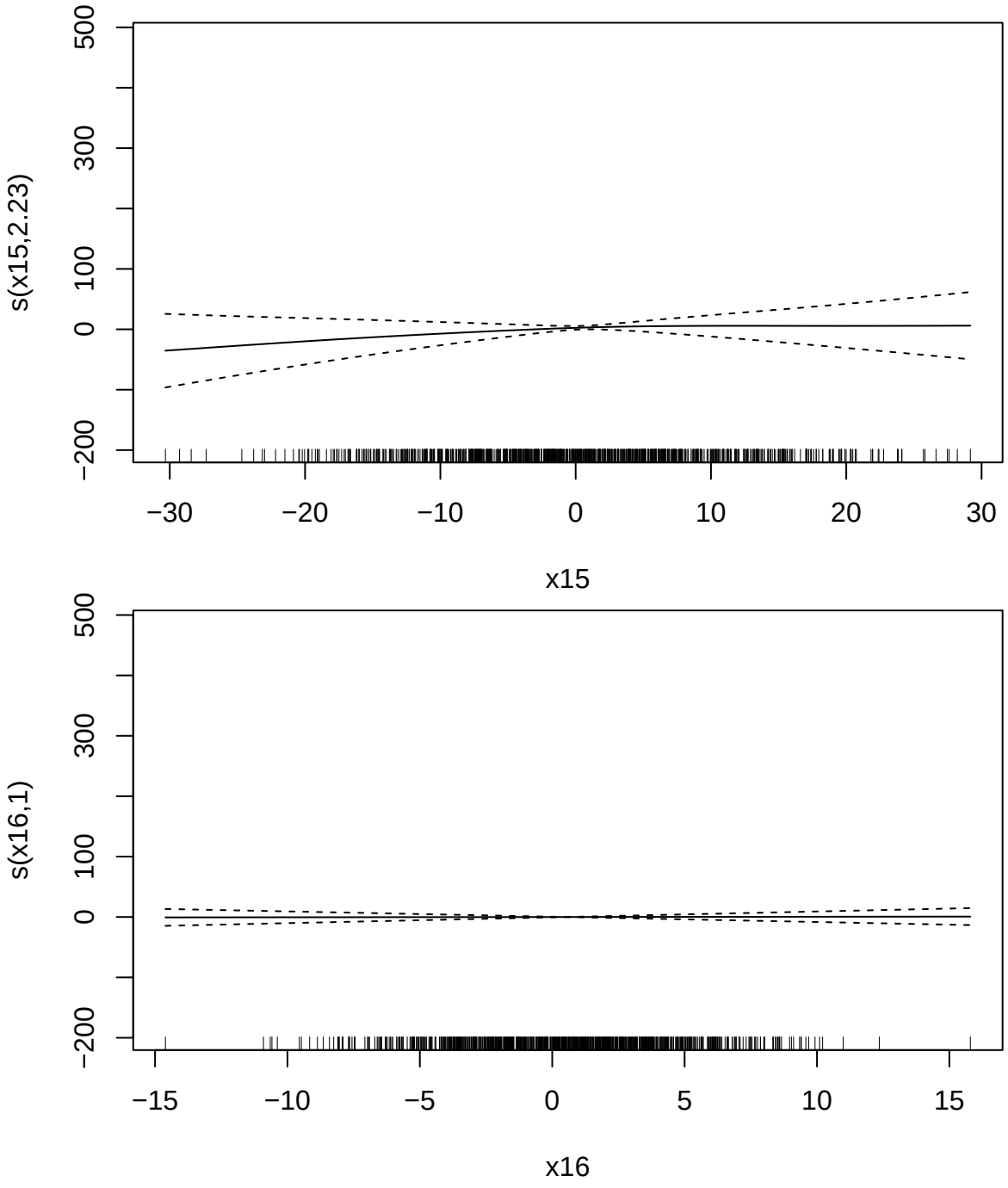

s(x9,1)
500
300
100
0
−200
0
1
2
3
4
x9
s(x10,1)
500
300
100
0
−200
−10
−8
−6
−4
−2
0
2
x10

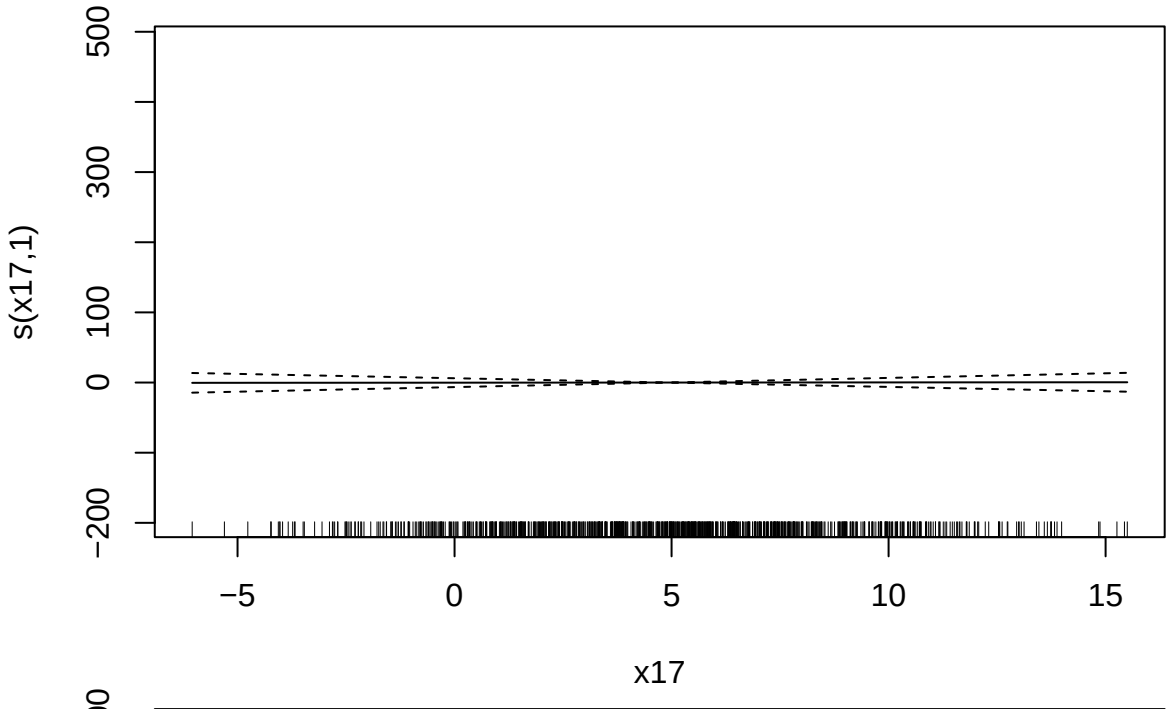

s(x11,1)
500
300
100
0
−200
−3
−2
−1
0
1
2
x11

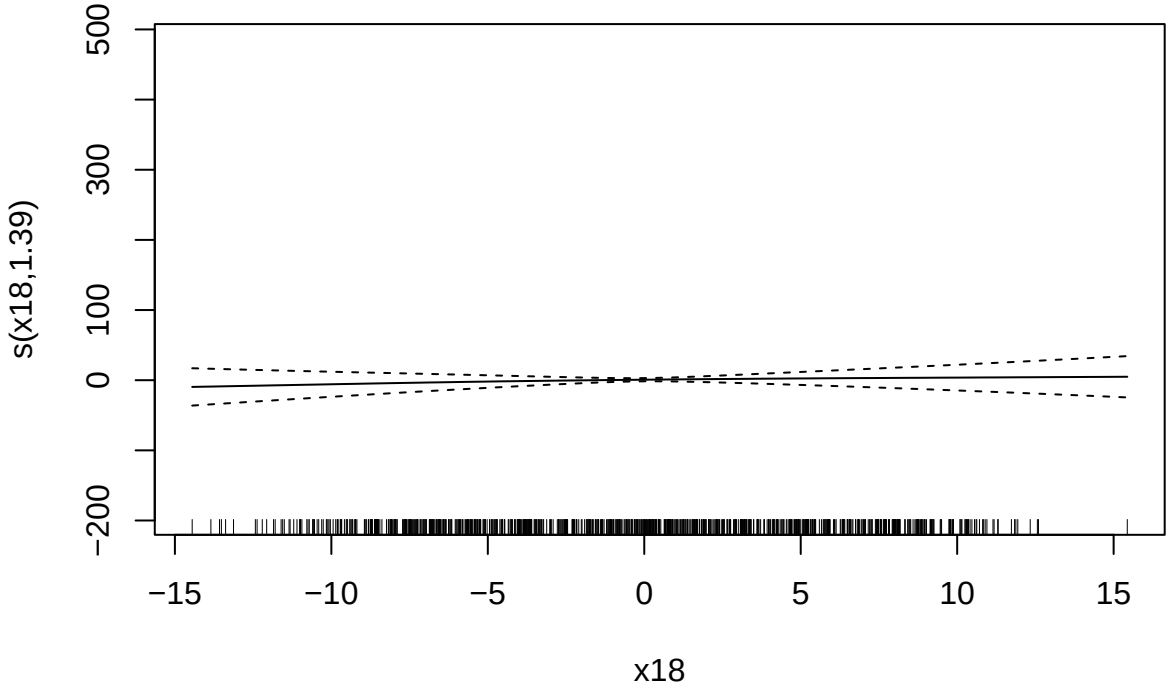

s(x12,6.31)
500
300
100
0
−200
−6
−4
−2
0
2
4
6
x12

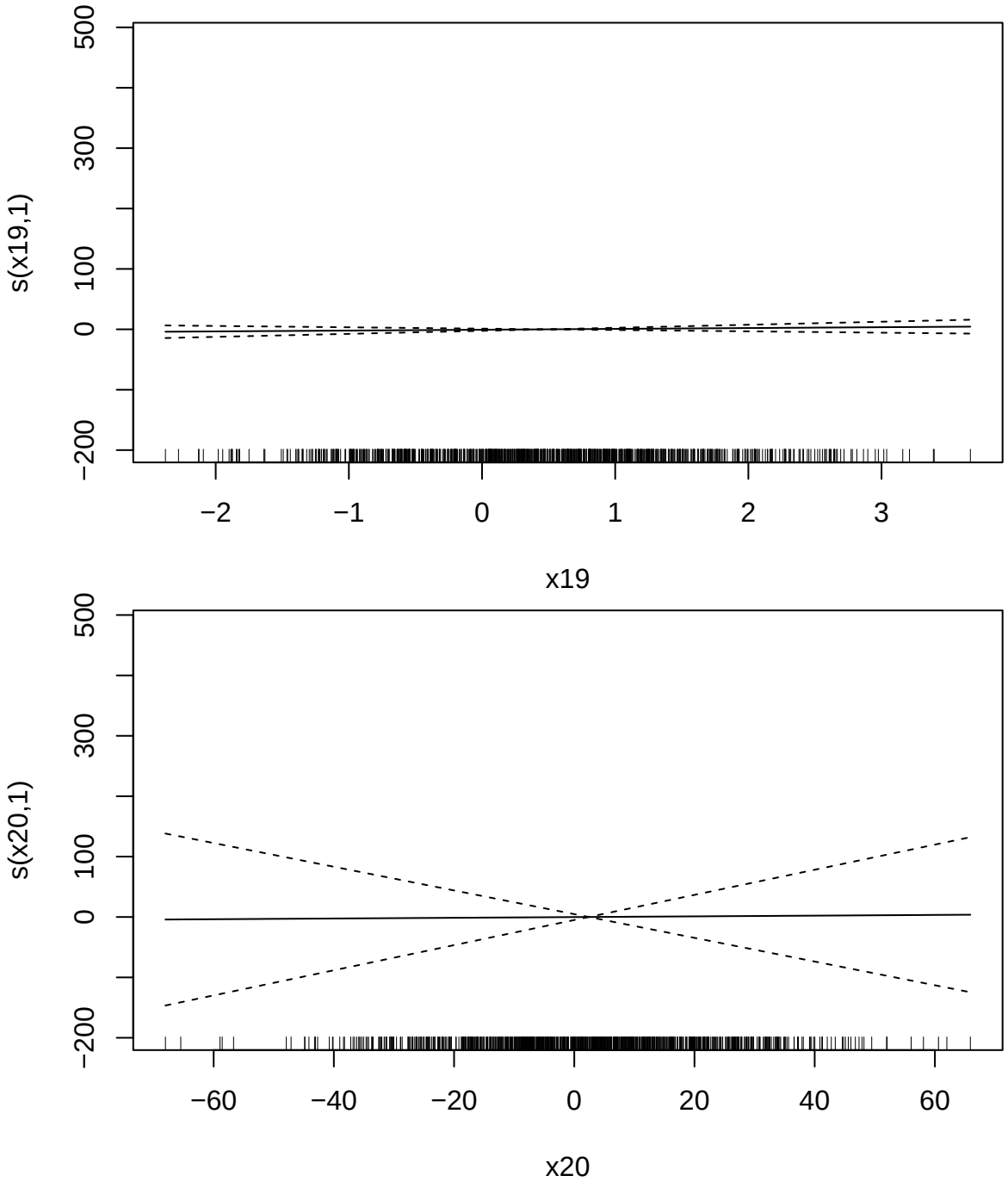

s(x13,3.83)
500
300
100
0
−200
−10
−5
0
5
10
x13

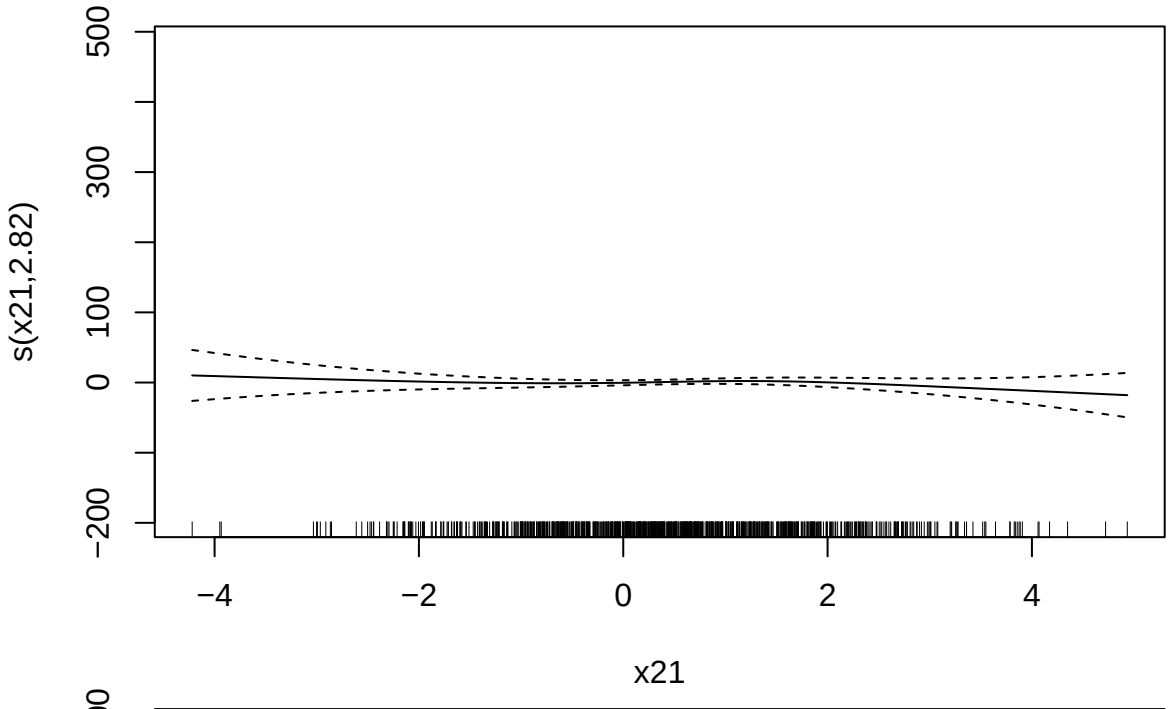

s(x14,1)
500
300
100
0
−200
−10
−5
0
5
10
x14

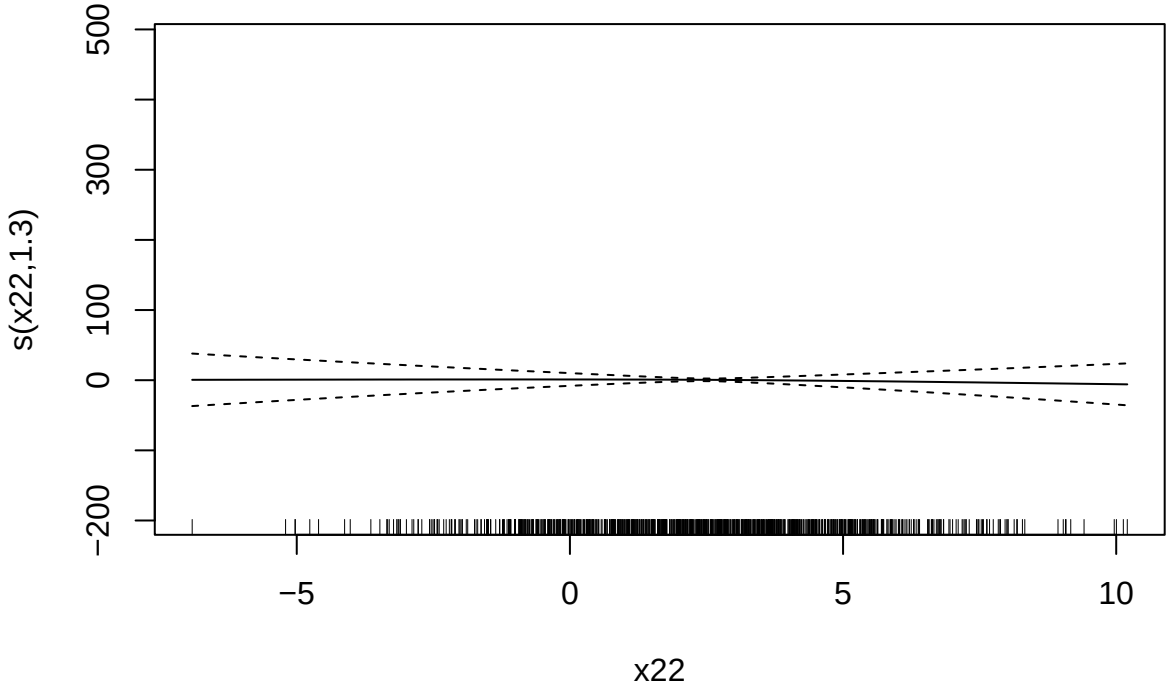

s(x15,2.23)
500
300
100
0
−200
−30
−20
−10
0
10
20
30
x15
s(x16,1)
500
300
100
0
−200
−15
−10
−5
0
5
10
15
x16

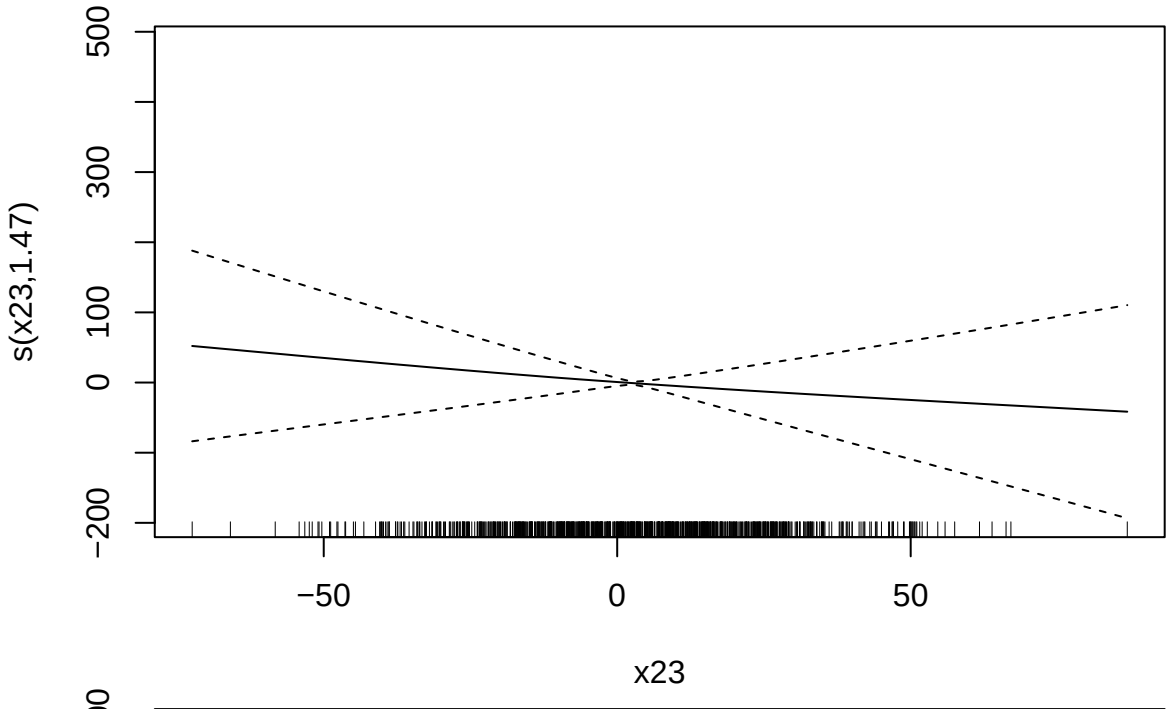

s(x17,1)
500
300
100
0
−200
−5
0
5
10
15
x17

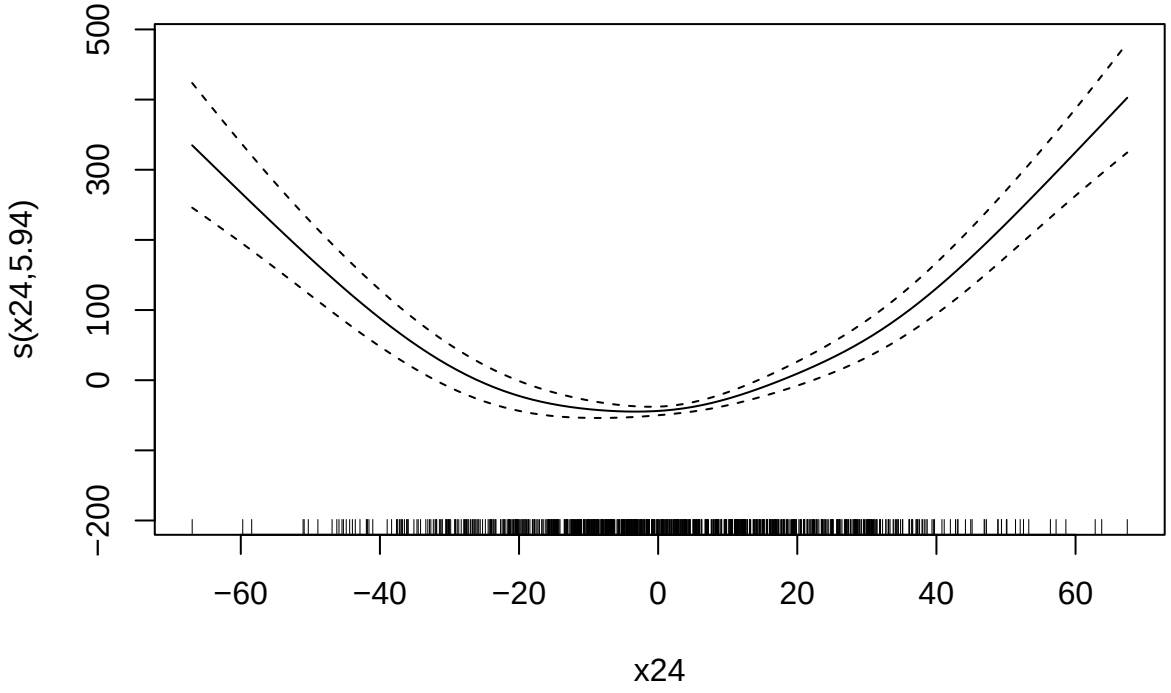

s(x18,1.39)
500
300
100
0
−200
−15
−10
−5
0
5
10
15
x18

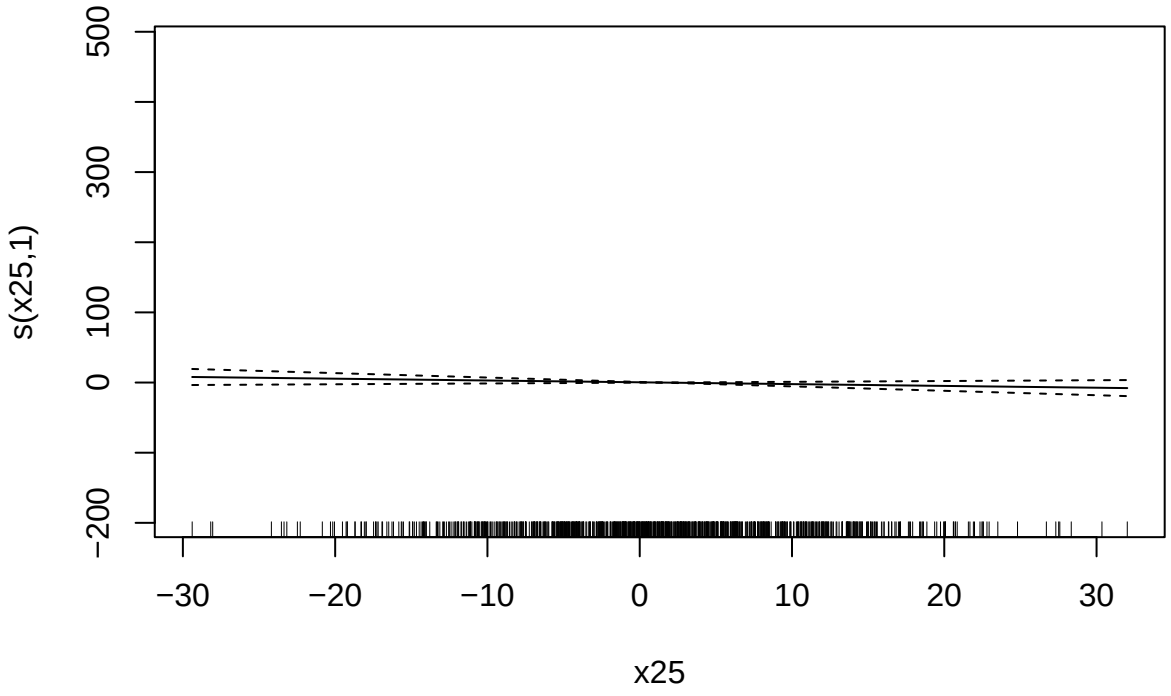

s(x19,1)
500
300
100
0
−200
−2
−1
0
1
2
3
x19
s(x20,1)
500
300
100
0
−200
−60
−40
−20
0
20
40
60
x20


s(x21,2.82)
500
300
100
0
−200
−4
−2
0
2
4
x21


s(x22,1.3)
500
300
100
0
−200
−5
0
5
10
x22

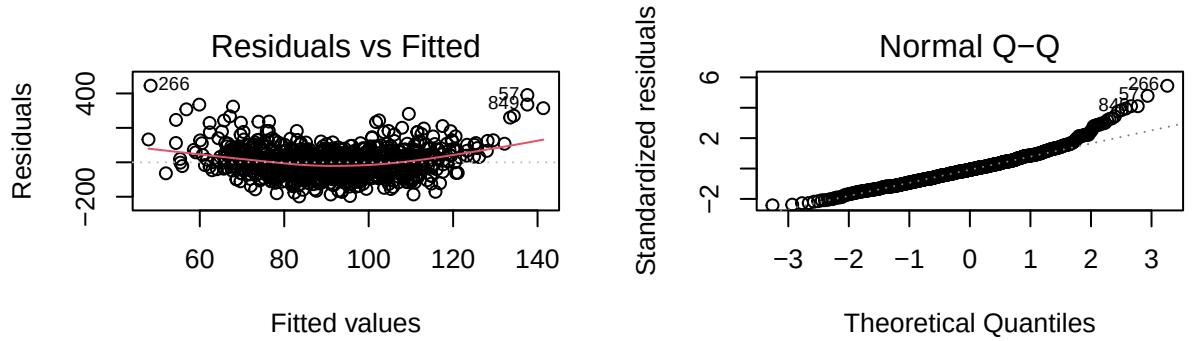

s(x23,1.47)
500
300
100
0
−200
−50
0
50
x23

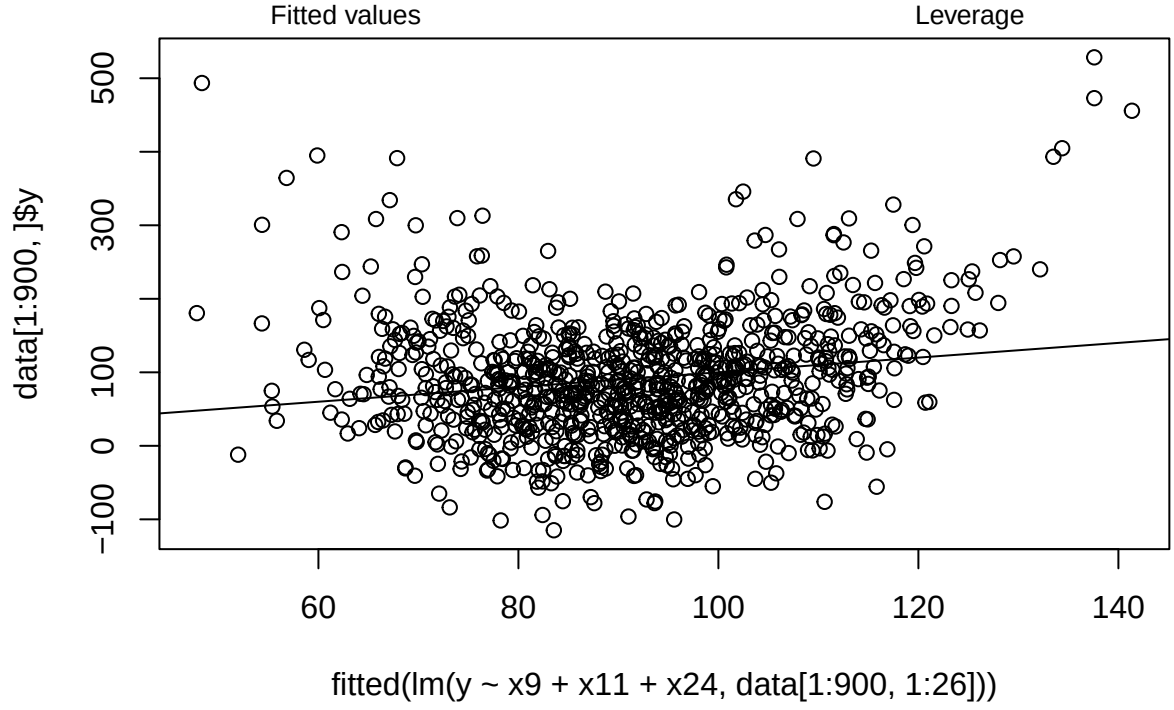

s(x24,5.94)
500
300
100
0
−200
−60
−40
−20
0
20
40
60
x24

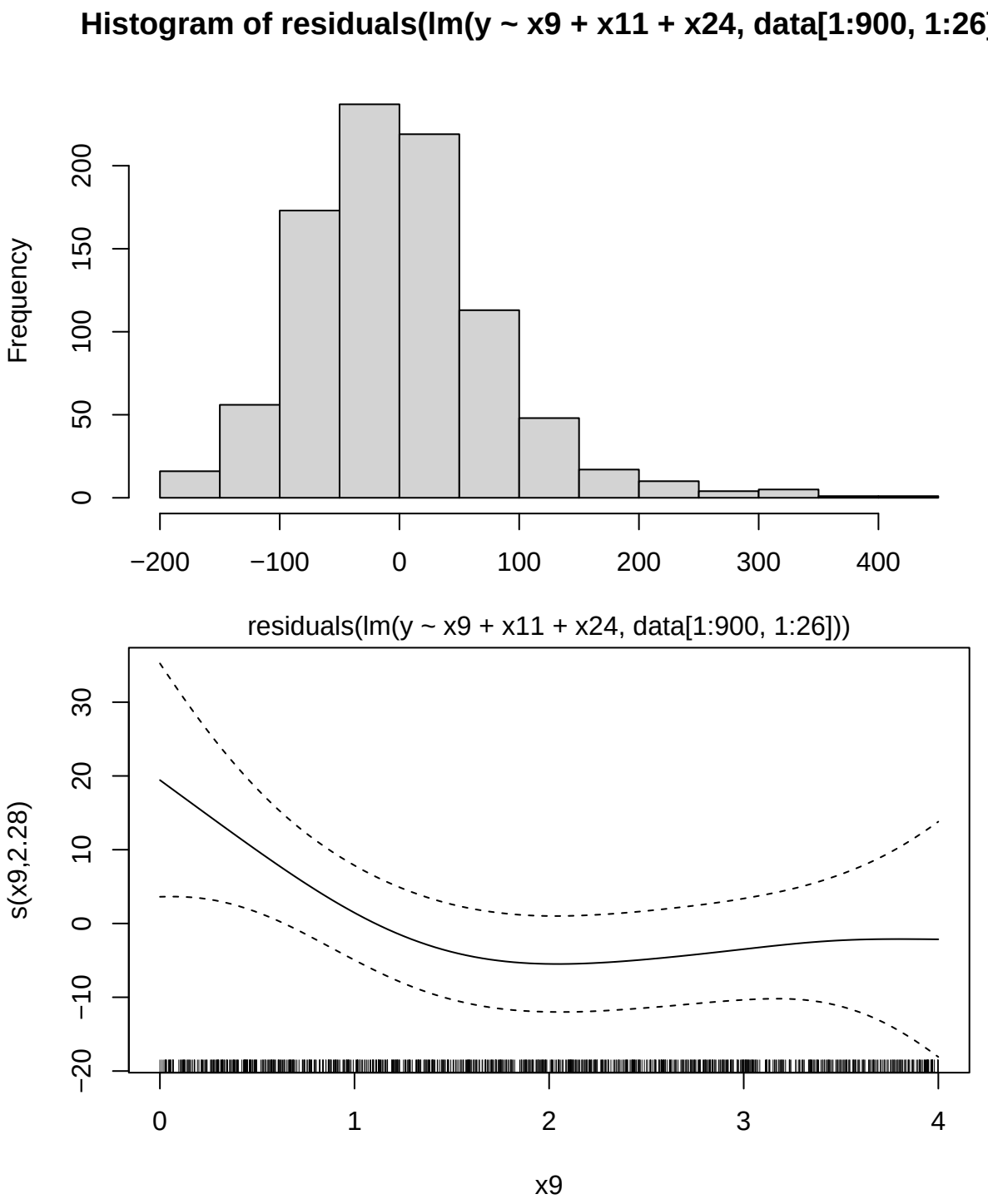


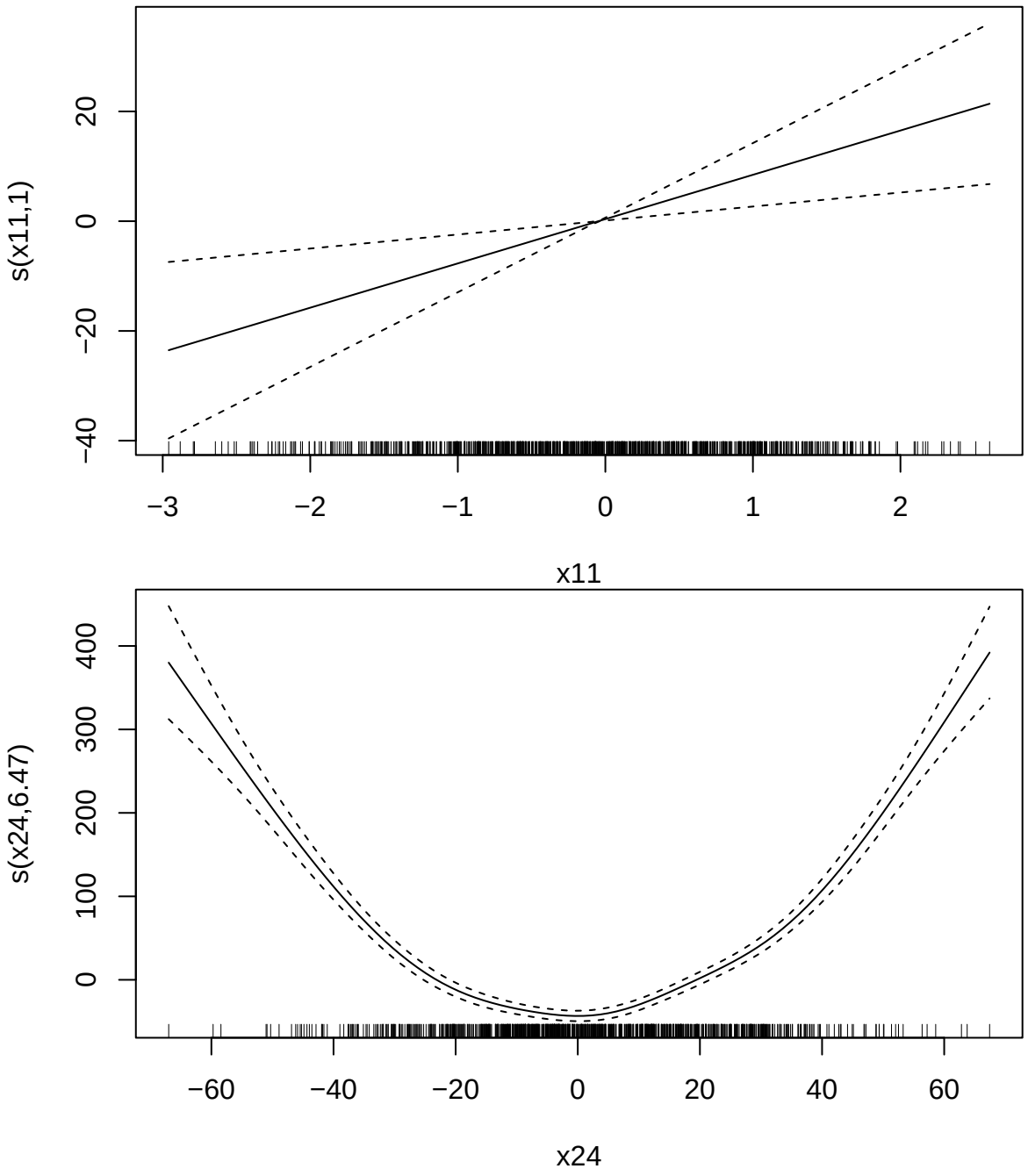

```
##
## Call:
## lm(formula = y ~ x9 + x11 + x24, data = data[1:900, 1:26])
##
## Coefficients:
## (Intercept)           x9          x11          x24
##     98.6098      -4.1152       8.7450       0.5691
```

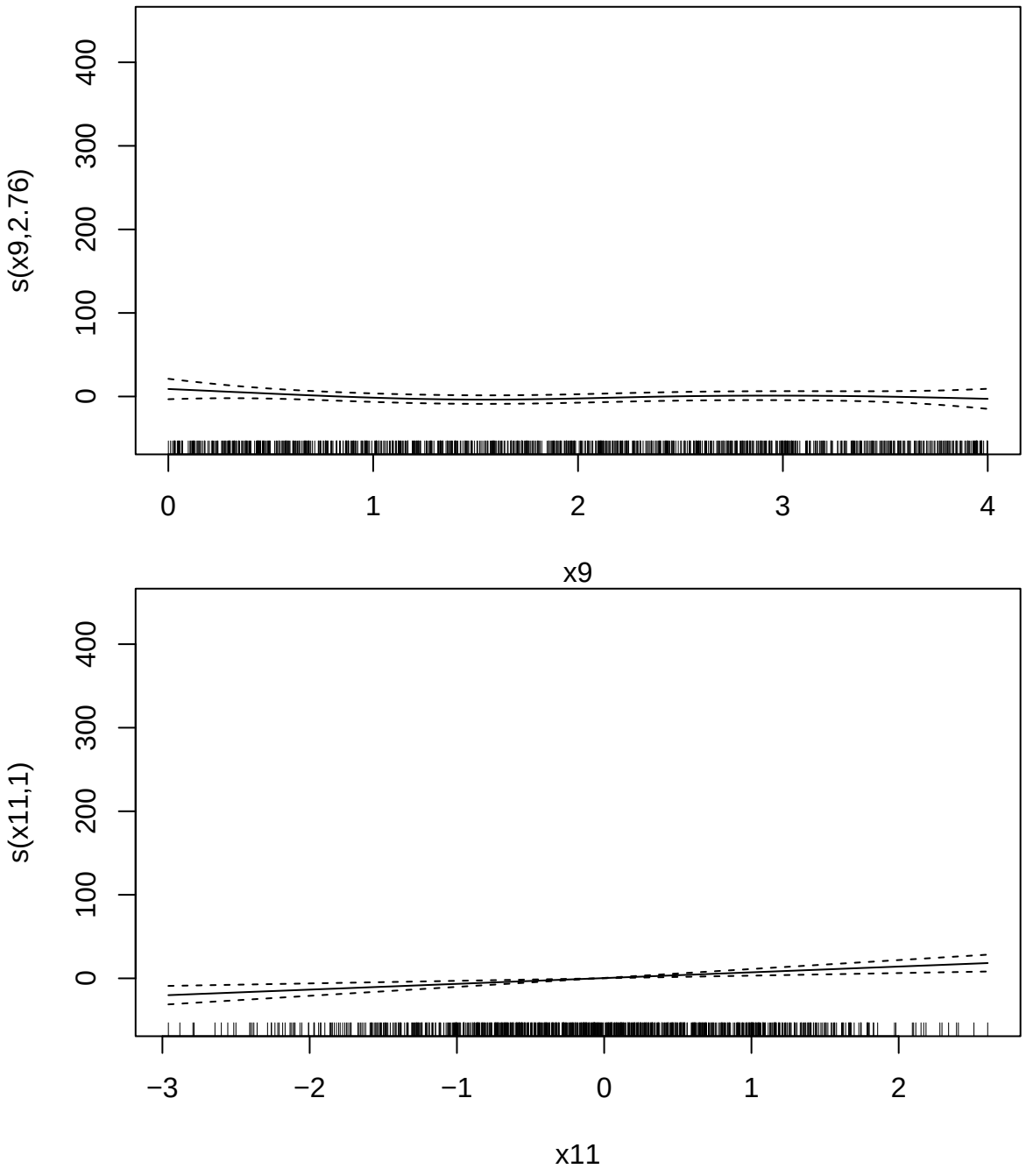
```
##
## Call:
## lm(formula = y ~ x9 + x11 + x24, data = data[1:900, 1:26])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -198.36  -54.31   -5.83   41.31  445.18
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  98.6098     5.4169  18.204  < 2e-16 ***
## x9           -4.1152     2.3985  -1.716  0.08656 .
## x11           8.7450     2.7412   3.190  0.00147 **
## x24           0.5691     0.1307   4.353 1.49e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 82.31 on 896 degrees of freedom
## Multiple R-squared:  0.03254,    Adjusted R-squared:  0.0293
## F-statistic: 10.05 on 3 and 896 DF,  p-value: 1.627e-06
```

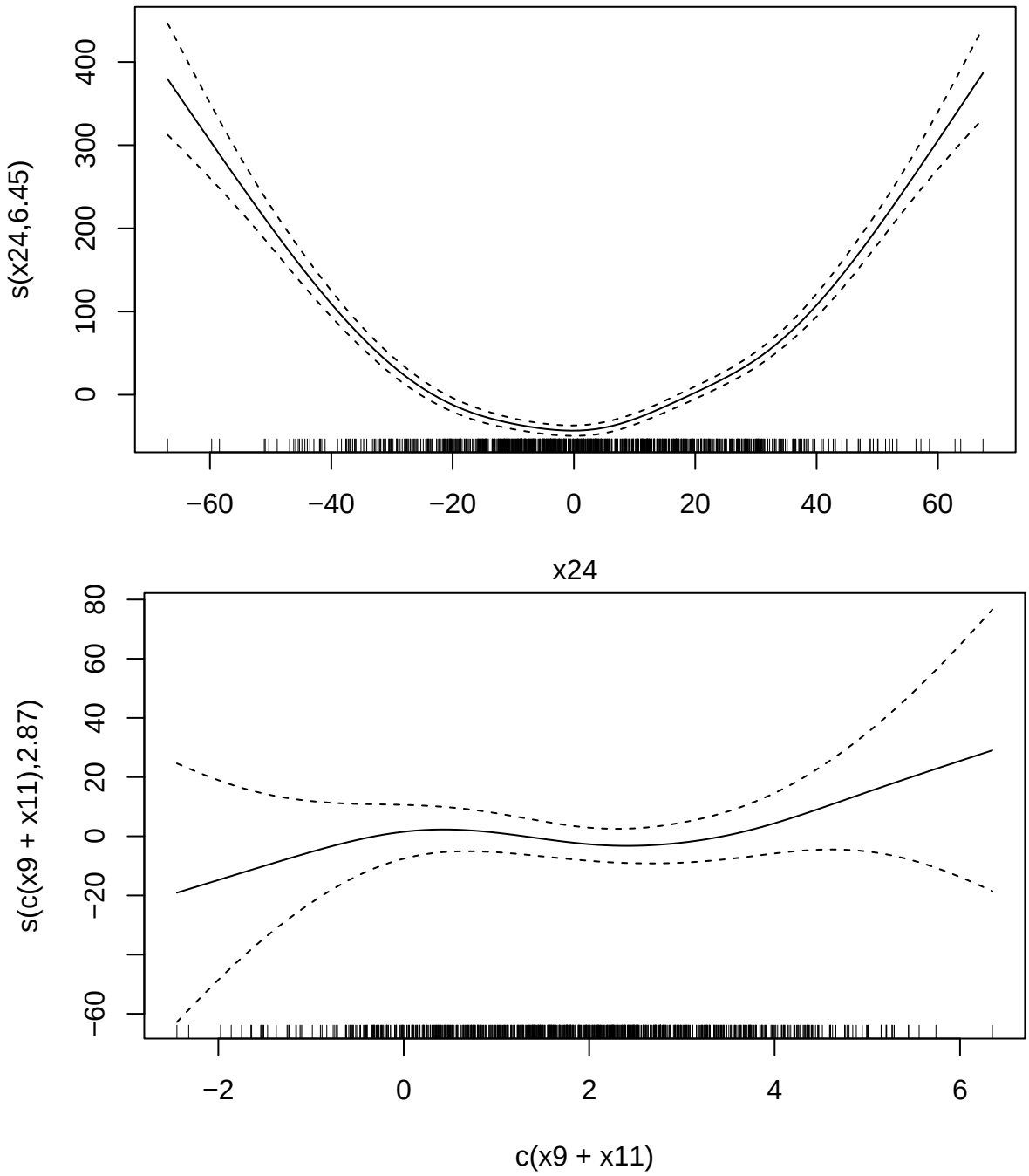

Residuals vs Fitted
Residuals
Fitted values
266
57
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Scale−Location
√|Standardized residuals|
Residuals vs Leverage
Cook's distance
Leverage

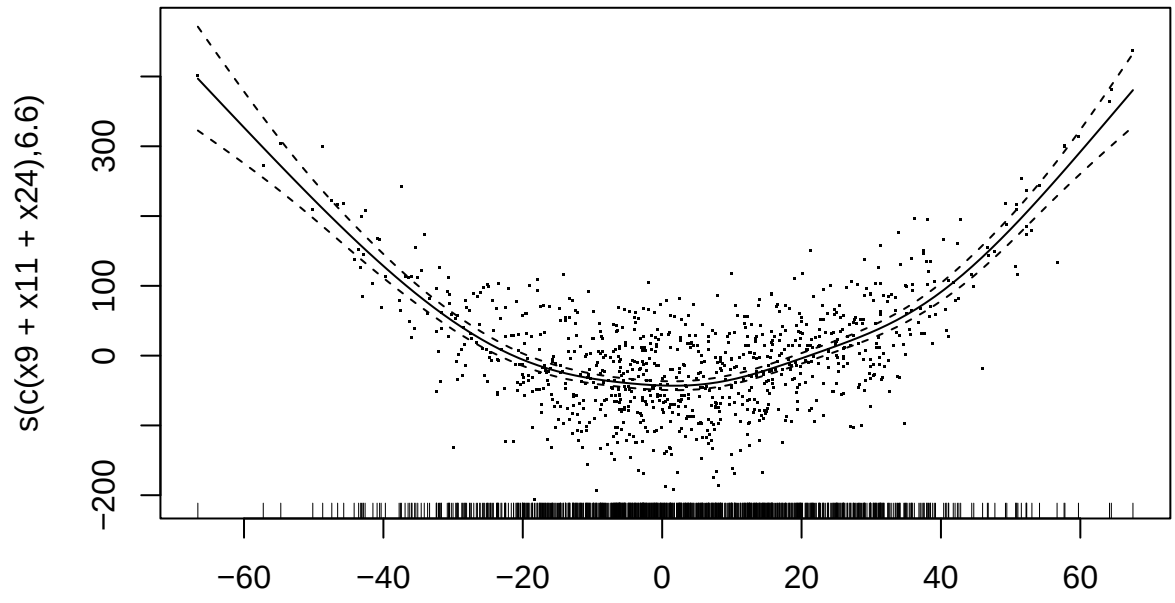

data[1:900,]$y
fitted(lm(y ~ x9 + x11 + x24, data[1:900, 1:26]))

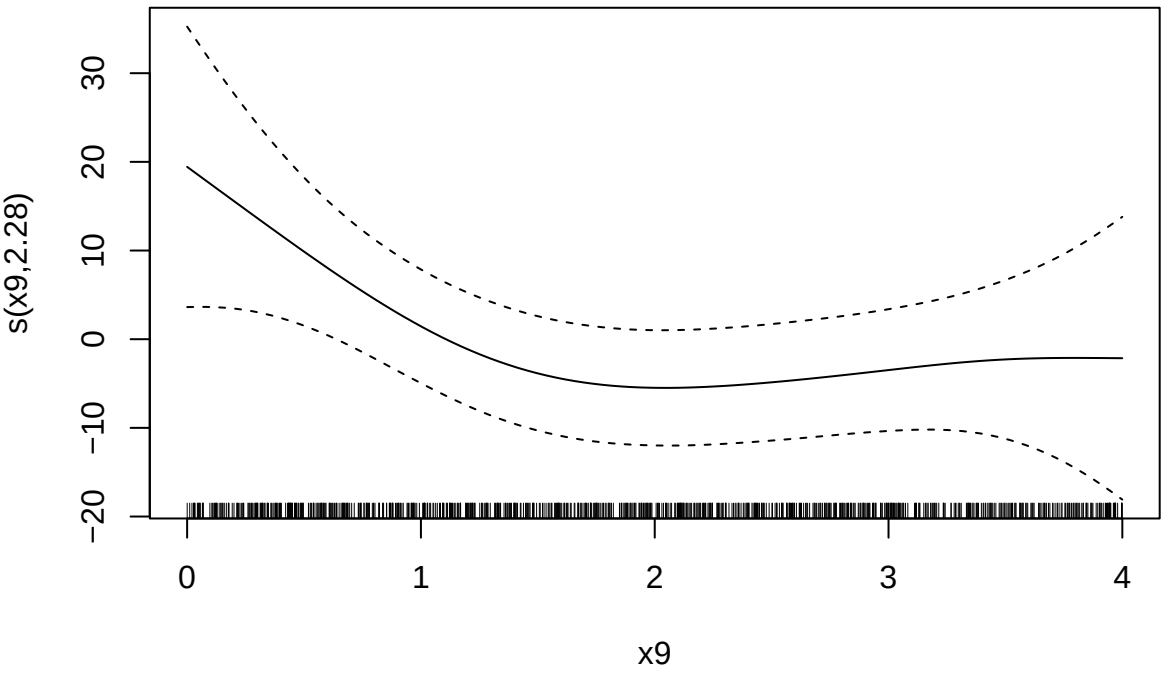

Histogram of residuals(lm(y ~ x9 + x11 + x24, data[1:900, 1:26]))
Frequency
0
50
100
150
200
−200
−100
0
100
200
300
400
residuals(lm(y ~ x9 + x11 + x24, data[1:900, 1:26]))
s(x9,2.28)
−20
−10
0
10
20
30
0
1
2
3
4
x9


s(x11,1)
20
0
−20
−40
−3
−2
−1
0
1
2
x11
s(x24,6.47)
400
300
200
100
0
−60
−40
−20
0
20
40
60
x24


s(x9,2.76)
0
100
200
300
400
0
1
2
3
4
x9
s(x11,1)
0
100
200
300
400
−3
−2
−1
0
1
2
x11


s(x24,6.45)
0
100
200
300
400
−60
−40
−20
0
20
40
60
x24
s(c(x9 + x11),2.87)
−60
−20
0
20
40
60
80
−2
0
2
4
6
c(x9 + x11)






```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## y ~ s(x9) + s(x11) + s(x24)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   91.430      1.896   48.21   <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##          edf Ref.df       F  p-value
## s(x9)  2.759  3.435   1.093 0.376168
## s(x11) 1.000  1.000  13.188 0.000298 ***
## s(x24) 6.445  7.635 131.177  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.536   Deviance explained = 54.1%
## GCV = 3277.8  Scale est. = 3237      n = 900
```