

46-926 Homework 4, Part 2

Pittsburgh Campus
Jingyi Guo

Data Preparation

Code:

```
library(fANCOVA)
trainset = read.table("/Users/apple/Desktop/ML/train.csv", sep=",", header=T)
fullrow = rep(FALSE,nrow(trainset))
for(i in 1:nrow(trainset))
{
  fullrow[i] = !any(is.na(trainset[i,29:147]))
}
varnames = c(paste("Ret_",2:120, sep = ""))
```

1. Use one-dimensional non-parametric regression to “smooth” the return series

Code:

```
nonpara_fitted = matrix(nrow=22390,ncol=119)
for(i in 1:119)
{
  nonpara_fitted[,i]=loess(Ret_PlusOne~get(varnames[i]),data=trainset,subset=fullrow,span=0.4,degree=1,parametric=FALSE)$fitted
}
nonpara_fitted=as.data.frame(nonpara_fitted)
```

2. Fit the model using "smoothed" predictors

Code:

```
y=trainset[208][fullrow,]
newvarnames=c(paste("V",1:119, sep = ""))
# Here V1, V2, ..., V119 refer to nonparametric fitted values of Ret_2, Ret_3, ...,
# Ret_120 respectively, and y refer to Ret_PlusOne.
newform=as.formula(paste("y ~ ",paste(newvarnames,collapse = "+")))
lmfitmodel = lm(newform,data=nonpara_fitted)
summary(lmfitmodel)
AIC(lmfitmodel)
```

Output:

Call:

```
lm(formula = newform, data = nonpara_fitted)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.23284	-0.01082	0.00028	0.01090	0.39900

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.0002540	0.0002653	-0.958	0.338269
V1	0.9298167	0.2044589	4.548	5.45e-06 ***
V2	0.3731273	0.2978211	1.253	0.210271
V3	-0.0637890	0.0553494	-1.152	0.249137
V4	1.0895377	0.2055477	5.301	1.16e-07 ***
V5	-0.1159562	0.0724919	-1.600	0.109707
V6	0.0954202	0.1193301	0.800	0.423932
V7	0.5022530	0.0708960	7.084	1.44e-12 ***
V8	0.2091418	0.1055549	1.981	0.047564 *
V9	0.8237882	0.1006296	8.186	2.84e-16 ***
V10	-0.0765249	0.0711112	-1.076	0.281881
V11	0.4288731	0.1149846	3.730	0.000192 ***
V12	0.2263392	0.0882228	2.566	0.010308 *
V13	-0.2811809	0.0952486	-2.952	0.003160 **
V14	0.1267429	0.0805229	1.574	0.115502
V15	0.1300630	0.3265796	0.398	0.690444
V16	0.7917560	0.1526396	5.187	2.15e-07 ***
V17	0.6975928	0.2278204	3.062	0.002201 **
V18	0.1424660	0.0955758	1.491	0.136079
V19	0.1122161	0.0766564	1.464	0.143239
V20	-0.1663566	0.1246043	-1.335	0.181864
V21	0.2496034	0.0563945	4.426	9.64e-06 ***
V22	0.0050947	0.0882281	0.058	0.953953
V23	0.4856522	0.1004256	4.836	1.33e-06 ***
V24	0.1722694	0.1916386	0.899	0.368700
V25	0.1435309	0.1469449	0.977	0.328695
V26	0.9131696	0.1768843	5.163	2.46e-07 ***
V27	0.4506322	0.1003471	4.491	7.13e-06 ***
V28	0.3735169	0.0560751	6.661	2.78e-11 ***
V29	-0.8154067	0.3074334	-2.652	0.008000 **
V30	0.2853259	0.0988782	2.886	0.003910 **
V31	0.3812982	0.0679409	5.612	2.02e-08 ***
V32	0.6418648	0.1318154	4.869	1.13e-06 ***
V33	0.6684854	0.0735708	9.086	< 2e-16 ***
V34	-0.1826991	0.0999241	-1.828	0.067506 .

V35	-0.3291876	0.0980831	-3.356 0.000792 ***
V36	-0.2315886	0.1239515	-1.868 0.061722 .
V37	-0.3987680	0.1141289	-3.494 0.000477 ***
V38	-0.1425576	0.1199534	-1.188 0.234672
V39	0.0657038	0.1034213	0.635 0.525238
V40	0.0425997	0.0711716	0.599 0.549480
V41	0.1004919	0.0658956	1.525 0.127269
V42	-0.4115026	0.1328287	-3.098 0.001951 **
V43	0.1905835	0.0753473	2.529 0.011433 *
V44	0.7543260	0.2043617	3.691 0.000224 ***
V45	0.4924198	0.1756599	2.803 0.005063 **
V46	0.5156130	0.1791975	2.877 0.004014 **
V47	0.5374007	0.1676146	3.206 0.001347 **
V48	0.1325242	0.1326042	0.999 0.317614
V49	0.5699168	0.0802927	7.098 1.30e-12 ***
V50	0.6742420	0.1679622	4.014 5.98e-05 ***
V51	-0.1859276	0.1431618	-1.299 0.194052
V52	0.3858105	0.1432298	2.694 0.007073 **
V53	-0.0066383	0.1121426	-0.059 0.952797
V54	0.5166031	0.1269344	4.070 4.72e-05 ***
V55	-0.8273625	0.4071545	-2.032 0.042159 *
V56	0.3223407	0.0690479	4.668 3.05e-06 ***
V57	0.1553120	0.2308305	0.673 0.501056
V58	-0.2783785	0.1348446	-2.064 0.038988 *
V59	0.4854206	0.1632931	2.973 0.002955 **
V60	0.5555596	0.1316198	4.221 2.44e-05 ***
V61	0.2414256	0.1748575	1.381 0.167385
V62	-1.1840743	0.2445310	-4.842 1.29e-06 ***
V63	0.0512757	0.1834048	0.280 0.779805
V64	0.0221123	0.1000767	0.221 0.825130
V65	0.4719641	0.0701220	6.731 1.73e-11 ***
V66	-0.2217743	0.2807641	-0.790 0.429597
V67	-0.2413923	0.1564272	-1.543 0.122806
V68	-0.3777451	0.1689460	-2.236 0.025369 *
V69	0.1052814	0.1420759	0.741 0.458688
V70	0.4779006	0.0685453	6.972 3.21e-12 ***
V71	0.2497755	0.0644411	3.876 0.000106 ***
V72	0.8881596	0.1419708	6.256 4.02e-10 ***
V73	0.5298977	0.1050957	5.042 4.64e-07 ***
V74	0.3197004	0.0878060	3.641 0.000272 ***
V75	0.4411413	0.0821725	5.368 8.02e-08 ***
V76	0.0495647	0.0899030	0.551 0.581425

V77	-0.0190461	0.1665091	-0.114 0.908934
V78	0.0753497	0.0936882	0.804 0.421255
V79	0.6164210	0.2648111	2.328 0.019933 *
V80	0.1909797	0.1016017	1.880 0.060163 .
V81	0.3374828	0.1685464	2.002 0.045263 *
V82	-0.3641121	0.0805281	-4.522 6.17e-06 ***
V83	0.1315365	0.0736195	1.787 0.073998 .
V84	0.4288398	0.1958738	2.189 0.028580 *
V85	0.4574666	0.0591138	7.739 1.05e-14 ***
V86	0.5627835	0.1323514	4.252 2.13e-05 ***
V87	0.2751809	0.2008193	1.370 0.170610
V88	-0.1285386	0.0536240	-2.397 0.016536 *
V89	-0.0061668	0.1089424	-0.057 0.954859
V90	-0.4094736	0.1552179	-2.638 0.008344 **
V91	-0.6631479	0.2021689	-3.280 0.001039 **
V92	-0.0176753	0.2850546	-0.062 0.950558
V93	0.1729951	0.1300056	1.331 0.183310
V94	-0.0786303	0.0761417	-1.033 0.301763
V95	0.6118097	0.1408329	4.344 1.40e-05 ***
V96	0.0087213	0.0950050	0.092 0.926859
V97	0.1517396	0.1270277	1.195 0.232280
V98	0.5705608	0.1255815	4.543 5.57e-06 ***
V99	-0.5793196	0.2090606	-2.771 0.005592 **
V100	0.4532658	0.0694832	6.523 7.02e-11 ***
V101	0.1465775	0.0840589	1.744 0.081217 .
V102	0.1868008	0.0654657	2.853 0.004329 **
V103	-0.5129193	0.1443154	-3.554 0.000380 ***
V104	0.5217671	0.0731874	7.129 1.04e-12 ***
V105	-0.2370622	0.0594651	-3.987 6.72e-05 ***
V106	0.8713803	0.0794853	10.963 < 2e-16 ***
V107	-0.1343314	0.2695415	-0.498 0.618228
V108	0.3233015	0.1595474	2.026 0.042739 *
V109	0.5676722	0.0930791	6.099 1.09e-09 ***
V110	0.2657275	0.1221216	2.176 0.029571 *
V111	-0.2228946	0.1720056	-1.296 0.195038
V112	0.4422194	0.0663586	6.664 2.73e-11 ***
V113	-0.2342920	0.0611622	-3.831 0.000128 ***
V114	1.3153277	0.1624194	8.098 5.85e-16 ***
V115	0.3730931	0.0755167	4.941 7.85e-07 ***
V116	-0.0324246	0.0702570	-0.462 0.644434
V117	-0.3397602	0.0965899	-3.518 0.000436 ***
V118	0.3384595	0.1028835	3.290 0.001004 **

V119 0.1241721 0.1109734 1.119 0.263179

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

Residual standard error: 0.02501 on 22270 degrees of freedom

Multiple R-squared: 0.1784, Adjusted R-squared: 0.174

F-statistic: 40.64 on 119 and 22270 DF, p-value: < 2.2e-16

Code:

AIC(lmfitmodel)

Output:

[1] -101506.5