Applied Exercise 11

Team 13 - Xander Giarracco, Xuanxiong Zhen, Hanzheng Li, Henry Tazewell

2/16/2021

Applied Exercise 11

GAMs are generally fit using a backfitting approach. This exercise explores the idea behind backfitting by approximating multiple linear regression using simple linear regression in an iterative process.

Suppose that we would like to perform multiple linear regression, but we do not have software to do so. Instead, we only have software to perform simple linear regression. Therefore, we take the following iterative approach: we repeatedly hold all but one coefficient estimate fixed at its current value, and update only that coefficient estimate using a simple linear regression. The process is continued until convergence—that is, until the coefficient estimates stop changing.

Generate a response Y and two predictors X1 and X2

```
y<-Hitters$Runs
x1<-Hitters$Hits
x2<-Hitters$AtBat
```

Here is the model we are trying to fit:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

Initialize β_1 to take on a value of your choice, in this example we will arbitrarily choose 1.

```
beta1<-1
```

We then rearrange the model such that β_0 and β_2 are held constant:

$$Y - \hat{\beta}_1 X_1 = \beta_0 + \beta_2 X_2 + \epsilon$$

Keeping β_1 fixed, fit the model:

```
a<-y-beta1*x1
beta2<-lm(a~x2)$coef[2]
```

Using the same idea as above, we then rearrange the model such that β_0 and β_1 are held constant:

$$Y - \hat{\beta}_2 X_2 = \beta_0 + \beta_1 X_1 + \epsilon$$

Keeping β_2 fixed, fit the model:

```r
a<-y-beta2*x2
beta1<-lm(a~x1)$coef[2]
```

Write a for loop to repeat the process 100 times. Report the estimates of β_0, β_1, and β_2 at each iteration of the for loop. We can view the print out below:

```r
betas<-data.frame(beta0=numeric(),beta1=numeric(),beta2=numeric())
x<-1:100
for(i in x) {
  a<-y-beta1*x1
  beta2<-lm(a~x2)$coef[2]

  a<-y-beta2*x2
  beta1<-lm(a~x1)$coef[2]

  beta0<-lm(a~x1)$coef[1]

  betas<-rbind(betas, data.frame(beta0=beta0,beta1=beta1,beta2=beta2))
  print(betas[i,])
}
```

```
##                beta0     beta1      beta2
## (Intercept) 6.031118 0.9194591 -0.1260325
##                 beta0     beta1      beta2
## (Intercept)1 5.368616 0.8829585 -0.1146131
##                 beta0     beta1      beta2
## (Intercept)2 4.747914 0.8487609 -0.1039143
##                 beta0    beta1       beta2
## (Intercept)3 4.166375 0.816721 -0.09389043
##                 beta0     beta1       beta2
## (Intercept)4 3.621528 0.7867026 -0.08449904
##                 beta0     beta1       beta2
## (Intercept)5 3.111058 0.7585782 -0.07570019
##                 beta0     beta1      beta2
## (Intercept)6 2.632796 0.7322283 -0.0674565
##                 beta0     beta1       beta2
## (Intercept)7 2.184709 0.7075409 -0.05973294
##                 beta0     beta1      beta2
## (Intercept)8 1.764894 0.6844112 -0.0524967
##                 beta0     beta1       beta2
## (Intercept)9 1.371567 0.6627408 -0.04571702
##                  beta0     beta1       beta2
## (Intercept)10 1.003057 0.6424378 -0.03936511
##                   beta0     beta1       beta2
## (Intercept)11 0.6577978 0.6234157 -0.03341396
##                   beta0     beta1      beta2
## (Intercept)12 0.3343225 0.6055938 -0.0278383
##                    beta0     beta1       beta2
## (Intercept)13 0.03125671 0.5888964 -0.02261443
##                    beta0     beta1       beta2
## (Intercept)14 -0.2526873 0.5732524 -0.01772015
##                    beta0     beta1       beta2
## (Intercept)15 -0.5187161 0.5585956 -0.01313468
```

```
##                    beta0     beta1        beta2
## (Intercept)16 -0.7679599 0.5448635 -0.008838531
##                   beta0     beta1        beta2
## (Intercept)17 -1.001478 0.5319978 -0.004813441
##                   beta0     beta1        beta2
## (Intercept)18 -1.220262 0.5199438 -0.001042312
##                   beta0     beta1       beta2
## (Intercept)19 -1.425242 0.5086505 0.002490879
##                   beta0     beta1       beta2
## (Intercept)20 -1.617289 0.4980696 0.005801146
##                   beta0     beta1       beta2
## (Intercept)21 -1.797219 0.4881564 0.008902554
##                   beta0     beta1      beta2
## (Intercept)22 -1.965797 0.4788686 0.01180828
##                   beta0     beta1      beta2
## (Intercept)23 -2.123738 0.4701668 0.01453067
##                   beta0     beta1      beta2
## (Intercept)24 -2.271714 0.4620141 0.01708129
##                   beta0     beta1      beta2
## (Intercept)25 -2.410353 0.4543757 0.01947099
##                   beta0     beta1     beta2
## (Intercept)26 -2.540245 0.4472193 0.0217099
##                   beta0     beta1      beta2
## (Intercept)27 -2.661942 0.4405144 0.02380756
##                  beta0     beta1      beta2
## (Intercept)28 -2.77596 0.4342326 0.02577286
##                   beta0     beta1      beta2
## (Intercept)29 -2.882784 0.4283471 0.02761416
##                   beta0     beta1      beta2
## (Intercept)30 -2.982868 0.4228329 0.02933929
##                   beta0     beta1      beta2
## (Intercept)31 -3.076638 0.4176667 0.03095557
##                   beta0     beta1      beta2
## (Intercept)32 -3.164491 0.4128264 0.03246988
##                   beta0     beta1      beta2
## (Intercept)33 -3.246801 0.4082915 0.03388864
##                   beta0     beta1      beta2
## (Intercept)34 -3.323918 0.4040428 0.03521788
##                   beta0     beta1      beta2
## (Intercept)35 -3.396169 0.4000621 0.03646325
##                   beta0     beta1      beta2
## (Intercept)36 -3.463862 0.3963326 0.03763005
##                   beta0     beta1      beta2
## (Intercept)37 -3.527283 0.3928384 0.03872323
##                   beta0     beta1      beta2
## (Intercept)38 -3.586703 0.3895646 0.03974744
##                   beta0     beta1      beta2
## (Intercept)39 -3.642374 0.3864974 0.04070703
##                   beta0     beta1      beta2
## (Intercept)40 -3.694533 0.3836238 0.04160607
##                 beta0     beta1      beta2
## (Intercept)41 -3.7434 0.3809314 0.04244838
##                   beta0     beta1      beta2
## (Intercept)42 -3.789184 0.3784089 0.04323755
```

```
##                   beta0     beta1      beta2
## (Intercept)43 -3.83208 0.3760456 0.04397693
##                    beta0     beta1      beta2
## (Intercept)44 -3.872269 0.3738314 0.04466966
##                    beta0     beta1      beta2
## (Intercept)45 -3.909922 0.3717569 0.04531868
##                  beta0     beta1      beta2
## (Intercept)46 -3.9452 0.3698132 0.04592675
##                    beta0     beta1      beta2
## (Intercept)47 -3.978251 0.3679923 0.04649646
##                    beta0     beta1      beta2
## (Intercept)48 -4.009218 0.3662862 0.04703022
##                   beta0     beta1     beta2
## (Intercept)49 -4.03823 0.3646877 0.0475303
##                    beta0     beta1      beta2
## (Intercept)50 -4.065412 0.3631901 0.04799883
##                    beta0    beta1     beta2
## (Intercept)51 -4.090879 0.361787 0.0484378
##                   beta0     beta1      beta2
## (Intercept)52 -4.11474 0.3604724 0.04884907
##                    beta0     beta1      beta2
## (Intercept)53 -4.137094 0.3592408 0.04923439
##                    beta0     beta1     beta2
## (Intercept)54 -4.158039 0.3580869 0.0495954
##                    beta0     beta1      beta2
## (Intercept)55 -4.177661 0.3570057 0.04993364
##                    beta0     beta1      beta2
## (Intercept)56 -4.196046 0.3559928 0.05025053
##                    beta0     beta1      beta2
## (Intercept)57 -4.213271 0.3550438 0.05054743
##                    beta0     beta1     beta2
## (Intercept)58 -4.229409 0.3541547 0.0508256
##                    beta0     beta1      beta2
## (Intercept)59 -4.244529 0.3533217 0.05108621
##                    beta0     beta1      beta2
## (Intercept)60 -4.258694 0.3525412 0.05133038
##                    beta0   beta1      beta2
## (Intercept)61 -4.271966 0.35181 0.05155915
##                    beta0     beta1      beta2
## (Intercept)62 -4.284401 0.3511249 0.05177348
##                    beta0     beta1      beta2
## (Intercept)63 -4.296051 0.3504831 0.05197429
##                    beta0     beta1      beta2
## (Intercept)64 -4.306966 0.3498817 0.05216243
##                    beta0     beta1     beta2
## (Intercept)65 -4.317192 0.3493183 0.0523387
##                    beta0     beta1      beta2
## (Intercept)66 -4.326773 0.3487904 0.05250384
##                   beta0     beta1      beta2
## (Intercept)67 -4.33575 0.3482958 0.05265857
##                   beta0     beta1      beta2
## (Intercept)68 -4.34416 0.3478325 0.05280354
##                   beta0     beta1      beta2
## (Intercept)69 -4.35204 0.3473984 0.05293935
```

```
##                   beta0     beta1     beta2
## (Intercept)70 -4.359422 0.3469916 0.0530666
##                   beta0     beta1      beta2
## (Intercept)71 -4.366339 0.3466105 0.05318582
##                   beta0     beta1      beta2
## (Intercept)72 -4.372819 0.3462535 0.05329752
##                  beta0    beta1      beta2
## (Intercept)73 -4.37889 0.345919 0.05340217
##                   beta0     beta1      beta2
## (Intercept)74 -4.384579 0.3456056 0.05350022
##                   beta0    beta1      beta2
## (Intercept)75 -4.389908 0.345312 0.05359208
##                   beta0     beta1      beta2
## (Intercept)76 -4.394901 0.3450369 0.05367815
##                   beta0     beta1      beta2
## (Intercept)77 -4.399579 0.3447792 0.05375878
##                   beta0     beta1      beta2
## (Intercept)78 -4.403962 0.3445377 0.05383433
##                   beta0     beta1      beta2
## (Intercept)79 -4.408069 0.3443114 0.05390511
##                   beta0     beta1      beta2
## (Intercept)80 -4.411916 0.3440995 0.05397143
##                   beta0     beta1      beta2
## (Intercept)81 -4.415521 0.3439009 0.05403356
##                   beta0     beta1      beta2
## (Intercept)82 -4.418898 0.3437148 0.05409177
##                   beta0     beta1      beta2
## (Intercept)83 -4.422062 0.3435405 0.05414631
##                   beta0     beta1     beta2
## (Intercept)84 -4.425026 0.3433772 0.0541974
##                   beta0     beta1      beta2
## (Intercept)85 -4.427804 0.3432241 0.05424528
##                   beta0     beta1      beta2
## (Intercept)86 -4.430406 0.3430808 0.05429013
##                   beta0     beta1      beta2
## (Intercept)87 -4.432844 0.3429465 0.05433215
##                   beta0     beta1      beta2
## (Intercept)88 -4.435128 0.3428206 0.05437152
##                   beta0     beta1      beta2
## (Intercept)89 -4.437268 0.3427027 0.05440841
##                   beta0     beta1      beta2
## (Intercept)90 -4.439273 0.3425922 0.05444297
##                   beta0     beta1      beta2
## (Intercept)91 -4.441151 0.3424887 0.05447535
##                   beta0     beta1      beta2
## (Intercept)92 -4.442911 0.3423918 0.05450568
##                  beta0     beta1      beta2
## (Intercept)93 -4.44456 0.3423009 0.05453411
##                   beta0     beta1      beta2
## (Intercept)94 -4.446105 0.3422158 0.05456074
##                   beta0     beta1      beta2
## (Intercept)95 -4.447553 0.3421361 0.05458568
##                   beta0     beta1      beta2
## (Intercept)96 -4.448909 0.3420614 0.05460906
```

```
##                    beta0     beta1      beta2
## (Intercept)97 -4.450179 0.3419914 0.05463096
##                   beta0     beta1      beta2
## (Intercept)98 -4.45137 0.3419258 0.05465148
##                    beta0     beta1     beta2
## (Intercept)99 -4.452485 0.3418643 0.0546707
```

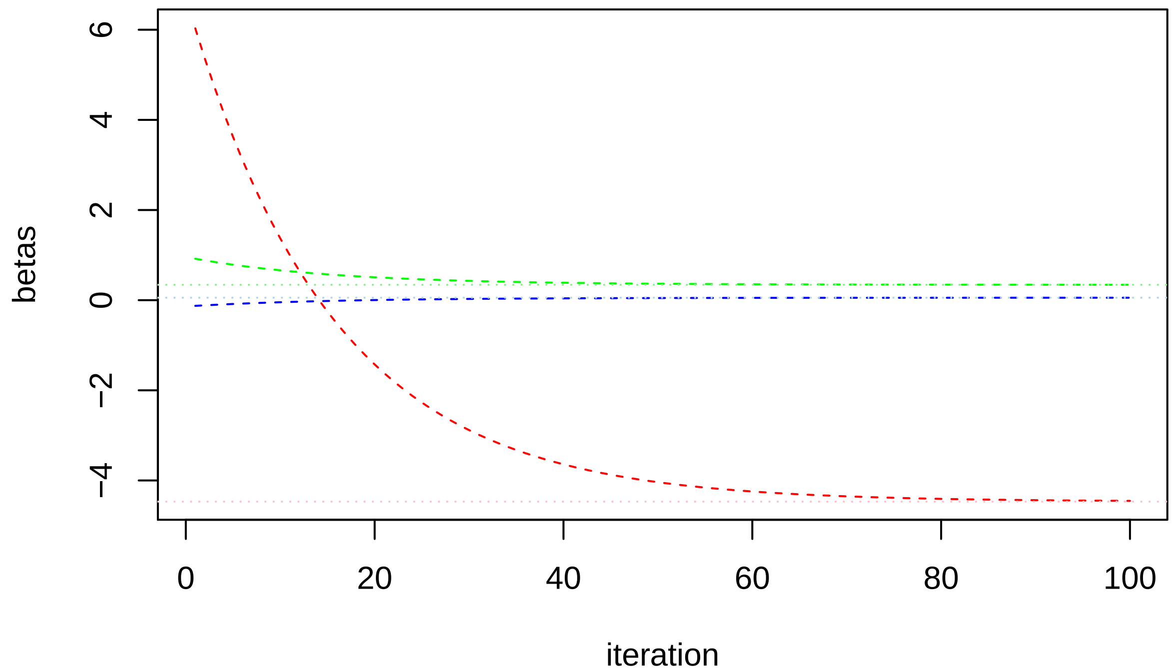
Plot the each β in a different color. Then create the proper multiple linear regression model using lm() and use abline() to plot the values of each new β in the same view.

```
plot(x,betas$beta0,type="l",col="red",lty=2,xlab='iteration',ylab='betas')
lines(x,betas$beta1,type="l",col="green",lty=2)
lines(x,betas$beta2,type="l",col="blue",lty=2)

mod<-lm(y~x1+x2)
summary(mod)
```

```
##
## Call:
## lm(formula = y ~ x1 + x2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -27.050  -5.872  -0.309   5.480  46.503
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.46905    1.55750  -2.869 0.004387 **
## x1           0.34095    0.04721   7.222 3.79e-12 ***
## x2           0.05496    0.01430   3.844 0.000146 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.87 on 319 degrees of freedom
## Multiple R-squared:  0.8571, Adjusted R-squared:  0.8562
## F-statistic: 956.3 on 2 and 319 DF,  p-value: < 2.2e-16
```

```
abline(mod$coef[1],0, col='pink',lty=3)
abline(mod$coef[2],0,col='lightgreen',lty=3)
abline(mod$coef[3],0,col='lightblue',lty=3)
```



On this data set, it took roughly 70 backfitting iterations to obtain a near perfect approximation for all three of the multiple regression coefficient estimates.