chapter\_2\_R\_exercise.R

RYU

Tue Nov 13 09:16:31 2018

library(lubridate)

##   
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':  
##   
## date

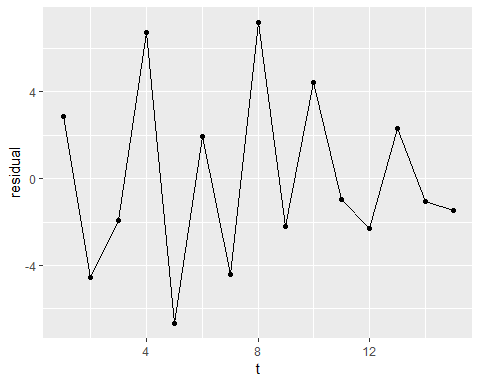
library(ggplot2)  
library(car)

## Loading required package: carData

library(forecast)  
  
# exercise 2.2  
z <- c(303,298,303,314,303,314,310,324,317,326,323,324,331,330,332)  
t <- 1:15  
df <- data.frame(t, z)  
reg <- lm(z~t, data=df)  
summary(reg)

##   
## Call:  
## lm(formula = z ~ t, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.664 -2.246 -1.071 2.579 7.200   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 297.7714 2.3164 128.547 < 2e-16 \*\*\*  
## t 2.3786 0.2548 9.336 3.97e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.263 on 13 degrees of freedom  
## Multiple R-squared: 0.8702, Adjusted R-squared: 0.8602   
## F-statistic: 87.16 on 1 and 13 DF, p-value: 3.969e-07

res <- data.frame(t, reg$residuals)  
colnames(res) <- c('t', 'residual')  
ggplot(data=res, aes(x=t,y=residual)) +   
 geom\_line() +  
 geom\_point()



new\_t <- 1:20  
new <- data.frame(t=new\_t)  
df <- data.frame(predict(reg, new, interval="prediction", level=0.95))  
df['t'] <- new\_t  
df['z'] <- c(z, rep('NA', 5))  
print(df)

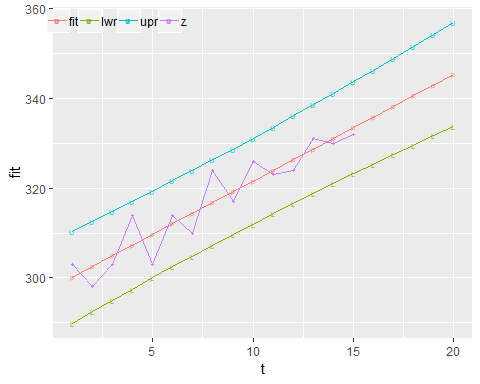
## fit lwr upr t z  
## 1 300.1500 289.8872 310.4128 1 303  
## 2 302.5286 292.4595 312.5976 2 298  
## 3 304.9071 295.0049 314.8093 3 303  
## 4 307.2857 297.5222 317.0493 4 314  
## 5 309.6643 300.0099 319.3186 5 303  
## 6 312.0429 302.4673 321.6184 6 314  
## 7 314.4214 304.8934 323.9494 7 310  
## 8 316.8000 307.2879 326.3121 8 324  
## 9 319.1786 309.6506 328.7066 9 317  
## 10 321.5571 311.9816 331.1327 10 326  
## 11 323.9357 314.2814 333.5901 11 323  
## 12 326.3143 316.5507 336.0778 12 324  
## 13 328.6929 318.7907 338.5951 13 331  
## 14 331.0714 321.0024 341.1405 14 330  
## 15 333.4500 323.1872 343.7128 15 332  
## 16 335.8286 325.3468 346.3104 16 NA  
## 17 338.2071 327.4825 348.9318 17 NA  
## 18 340.5857 329.5960 351.5755 18 NA  
## 19 342.9643 331.6888 354.2398 19 NA  
## 20 345.3429 333.7625 356.9232 20 NA

ggplot(data=df, aes(x=t)) +  
 geom\_line(aes(y=fit, colour='fit')) +  
 geom\_line(aes(y=lwr, colour='lwr')) +  
 geom\_line(aes(y=upr, colour='upr')) +  
 geom\_line(aes(x=t,y=as.numeric(z), colour='z')) +  
 geom\_point(aes(y=fit, colour='fit'), shape='p') +  
 geom\_point(aes(y=lwr, colour='lwr'), shape='L') +  
 geom\_point(aes(y=upr, colour='upr'), shape='U') +  
 geom\_point(aes(x=t,y=as.numeric(z), colour='z'), shape=20) +  
 theme(legend.position = c(0.15, 0.96),   
 legend.background=element\_rect(fill="transparent"),   
 legend.title=element\_blank(),  
 legend.direction = "horizontal")

## Warning in FUN(X[[i]], ...): 강제형변환에 의해 생성된 NA 입니다  
  
## Warning in FUN(X[[i]], ...): 강제형변환에 의해 생성된 NA 입니다

## Warning: Removed 5 rows containing missing values (geom\_path).

## Warning: Removed 5 rows containing missing values (geom\_point).



# exercise 2.9  
data <- read.csv('../timedata/catv.txt', sep='', header=FALSE)  
z <- na.omit(c(t(data)))  
t <- 1:length(z)  
df <- data.frame(t, z)  
reg <- lm(z~t, data=df)  
summary(reg)

##   
## Call:  
## lm(formula = z ~ t, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -7048466 -2080050 1364592 1849099 7087070   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -5305916 1560396 -3.40 0.00257 \*\*   
## t 2716876 109205 24.88 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3703000 on 22 degrees of freedom  
## Multiple R-squared: 0.9657, Adjusted R-squared: 0.9641   
## F-statistic: 618.9 on 1 and 22 DF, p-value: < 2.2e-16

new\_t <- 1:(length(z)+12)  
new <- data.frame(t=new\_t)  
df <- data.frame(predict(reg, new, interval="prediction", level=0.95))  
df['t'] <- new\_t  
df['z'] <- c(z, rep('NA', 12))  
print(df)

## fit lwr upr t z  
## 1 -2589039.9 -10848994.8 5670915 1 4498030  
## 2 127836.4 -8063526.9 8319200 2 5569810  
## 3 2844712.6 -5283792.3 10973218 3 6484380  
## 4 5561588.9 -2509925.1 13633103 4 7163340  
## 5 8278465.2 257949.8 16298981 5 8230310  
## 6 10995341.4 3019717.4 18970965 6 9196690  
## 7 13712217.7 5775274.0 21649161 7 10787970  
## 8 16429093.9 8524528.4 24333659 8 12168450  
## 9 19145970.2 11267403.0 27024537 9 13391910  
## 10 21862846.4 14003834.5 29721858 10 14814380  
## 11 24579722.7 16733774.7 32425671 11 17671490  
## 12 27296599.0 19457191.0 35136007 12 23219200  
## 13 30013475.2 22174067.3 37852883 13 29340570  
## 14 32730351.5 24884403.4 40576300 14 34113790  
## 15 35447227.7 27588215.8 43306240 15 37290870  
## 16 38164104.0 30285536.8 46042671 16 39872520  
## 17 40880980.2 32976414.7 48785546 17 42237140  
## 18 43597856.5 35660912.8 51534800 18 44970880  
## 19 46314732.8 38339108.7 54290357 19 48636520  
## 20 49031609.0 41011093.7 57052124 20 50897080  
## 21 51748485.3 43676971.3 59819999 21 53900000  
## 22 54465361.5 46336856.5 62593867 22 56000000  
## 23 57182237.8 48990874.5 65373601 23 57868170  
## 24 59899114.0 51639159.1 68159069 24 59397390  
## 25 62615990.3 54281851.9 70950129 25 NA  
## 26 65332866.5 56919100.8 73746632 26 NA  
## 27 68049742.8 59551058.9 76548427 27 NA  
## 28 70766619.1 62177883.0 79355355 28 NA  
## 29 73483495.3 64799732.9 82167258 29 NA  
## 30 76200371.6 67416770.0 84983973 30 NA  
## 31 78917247.8 70029156.5 87805339 31 NA  
## 32 81634124.1 72637054.5 90631194 32 NA  
## 33 84351000.3 75240624.9 93461376 33 NA  
## 34 87067876.6 77840027.3 96295726 34 NA  
## 35 89784752.9 80435418.7 99134087 35 NA  
## 36 92501629.1 83026953.5 101976305 36 NA

ggplot(data=df, aes(x=t)) +  
 geom\_line(aes(y=fit, colour='fit')) +  
 geom\_line(aes(y=lwr, colour='lwr')) +  
 geom\_line(aes(y=upr, colour='upr')) +  
 geom\_line(aes(x=t,y=as.numeric(z), colour='z')) +  
 geom\_point(aes(y=fit, colour='fit'), shape='p') +  
 geom\_point(aes(y=lwr, colour='lwr'), shape='L') +  
 geom\_point(aes(y=upr, colour='upr'), shape='U') +  
 geom\_point(aes(x=t,y=as.numeric(z), colour='z'), shape=20) +  
 theme(legend.position = c(0.15, 0.96),   
 legend.background=element\_rect(fill="transparent"),   
 legend.title=element\_blank(),  
 legend.direction = "horizontal")

## Warning in FUN(X[[i]], ...): 강제형변환에 의해 생성된 NA 입니다

## Warning in FUN(X[[i]], ...): 강제형변환에 의해 생성된 NA 입니다

## Warning: Removed 12 rows containing missing values (geom\_path).

## Warning: Removed 12 rows containing missing values (geom\_point).

