

# R\_final

문혜린(201611696)

2021 12 21

## 문제 풀이

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.5      v purrr 0.3.4
## v tibble 3.1.6       v dplyr 1.0.7
## v tidyr 1.1.4        v stringr 1.4.0
## v readr 2.1.1        v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()      masks stats::lag()
```

### 1번

(1)

```
epsilon = rnorm(100)
epsilon

## [1] -0.108533538 0.373059108 -0.324428677 -0.301112897 1.082839817
## [6] -0.255465382 -1.926763458 -1.035337054 0.418080314 2.061520798
## [11] 2.223599041 0.113848365 -1.318647421 0.714931105 -0.342528700
## [16] -1.187504304 1.047136970 -0.513963155 0.894740552 -0.180265682
## [21] 0.818372046 -2.023504444 -0.938393557 0.420596498 -0.061818222
## [26] -0.558652465 -1.540214043 -1.710169973 -0.697728082 -1.432152329
## [31] -1.225970388 0.595878423 0.068652580 2.345993279 0.330018729
## [36] 0.767804066 -0.245307726 -0.046334489 -0.516316169 0.244610397
## [41] -2.330476225 -0.825842186 -1.138368624 -1.345158009 -1.715808199
## [46] -0.817249551 0.006995816 -0.984932244 0.305529541 -0.119425132
## [51] 1.154057218 0.027551841 0.976935039 -1.321918735 0.262460878
## [56] -0.487921998 0.088976950 1.445933084 -0.543260003 -1.462751619
## [61] 0.883224050 -1.583185729 -0.363624550 0.785630212 -0.270534863
## [66] 0.660470347 -1.516871167 -1.483834060 0.007072152 -1.331410652
## [71] 0.757733700 -0.328654525 -1.401586380 -0.513041159 1.399825253
## [76] 0.624120532 0.654653332 -0.811091907 -1.070130444 0.857760317
## [81] 1.386585469 -0.267248841 -1.416224155 -0.287105803 0.451729724
## [86] -0.896979066 -0.913095565 0.825188757 -0.500034630 0.070018667
## [91] 2.711475194 -2.620166058 1.596539778 -0.300459126 0.167013433
## [96] 1.060923372 0.936176621 0.217025974 -0.433868020 -1.325693783
```

(2)

```
i = 1:1000  
t_i <- 2*pi*i/1000  
x1 = sin(t_i)  
x2 = cos(4*t_i)  
head(x1)
```

```
## [1] 0.006283144 0.012566040 0.018848440 0.025130095 0.031410759 0.037690183
```

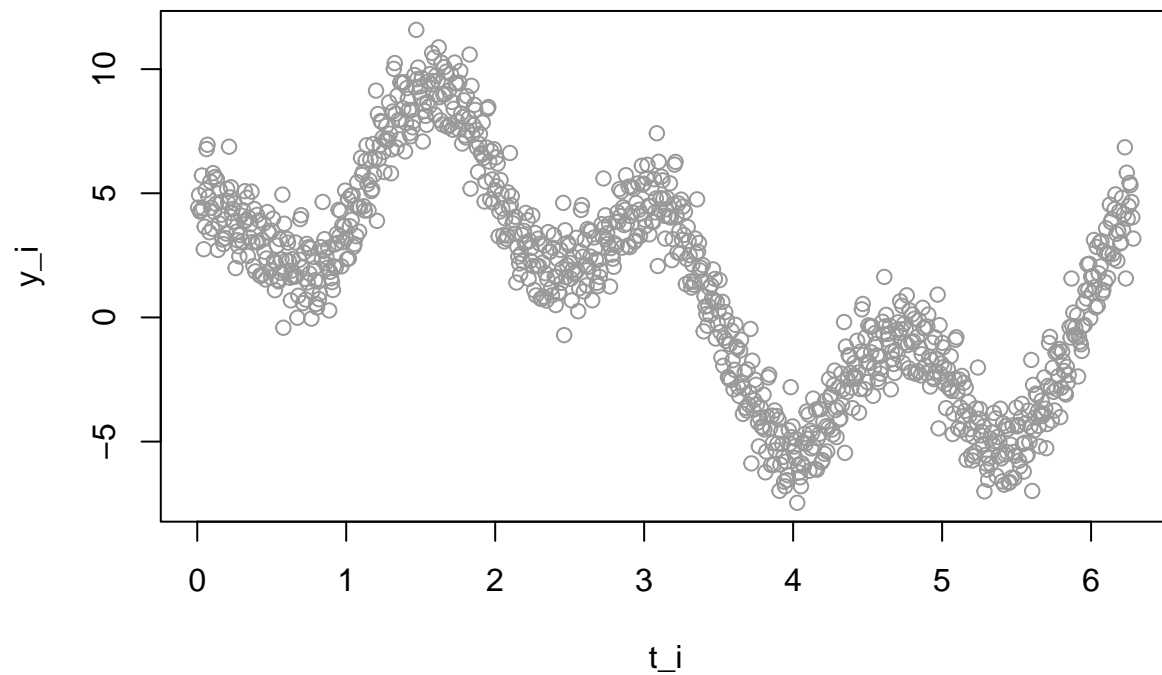
```
head(x2)
```

```
## [1] 0.9996842 0.9987370 0.9971589 0.9949510 0.9921147 0.9886517
```

(3)

```
y_i=1.5+5*x1+3*x2+epsilon
```

```
plot(t_i,y_i,col='gray60')
```



(4)

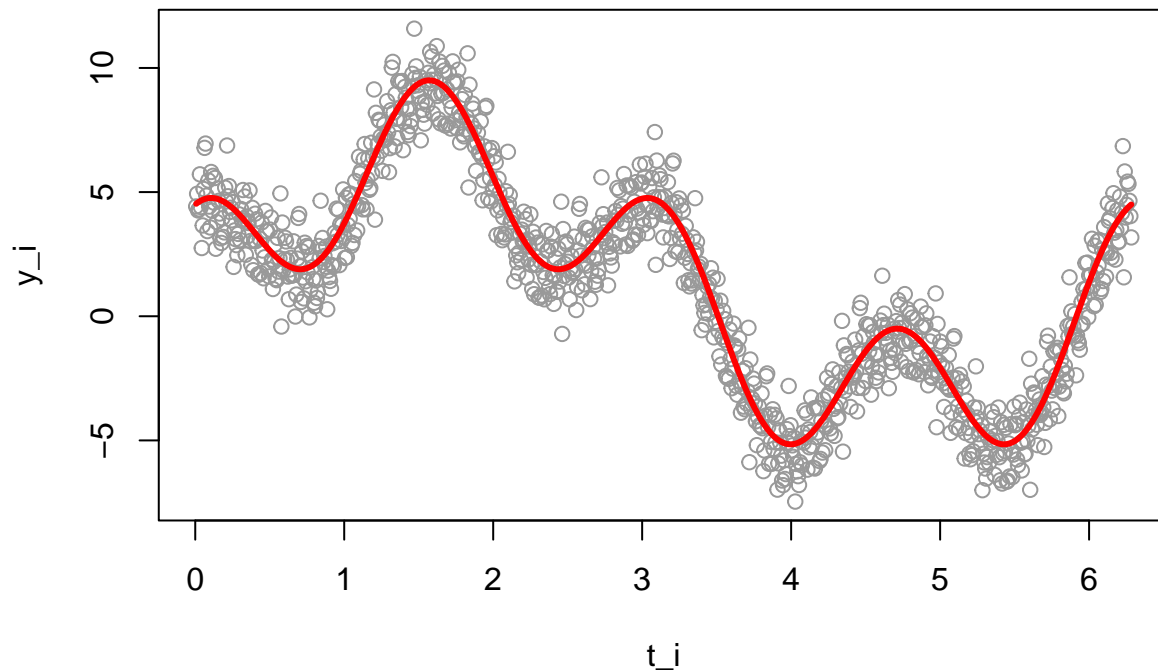
```
X=cbind(1,x1,x2)
```

(5)

```
B=rbind(1.5,5,3)
```

```
V <- X %*% B
```

```
plot(t_i,y_i,col='gray60')  
lines(t_i,V,col='red',lwd=3)
```



(6)

```
X_t=rbind(1,x1,x2)
y=cbind(y_i)
```

```
B_= solve(X_t %*% X) %*% X_t %*% y
B_
```

```
##      y_i
## 1.336242
## x1 5.000000
## x2 3.000000
```

```
B
```

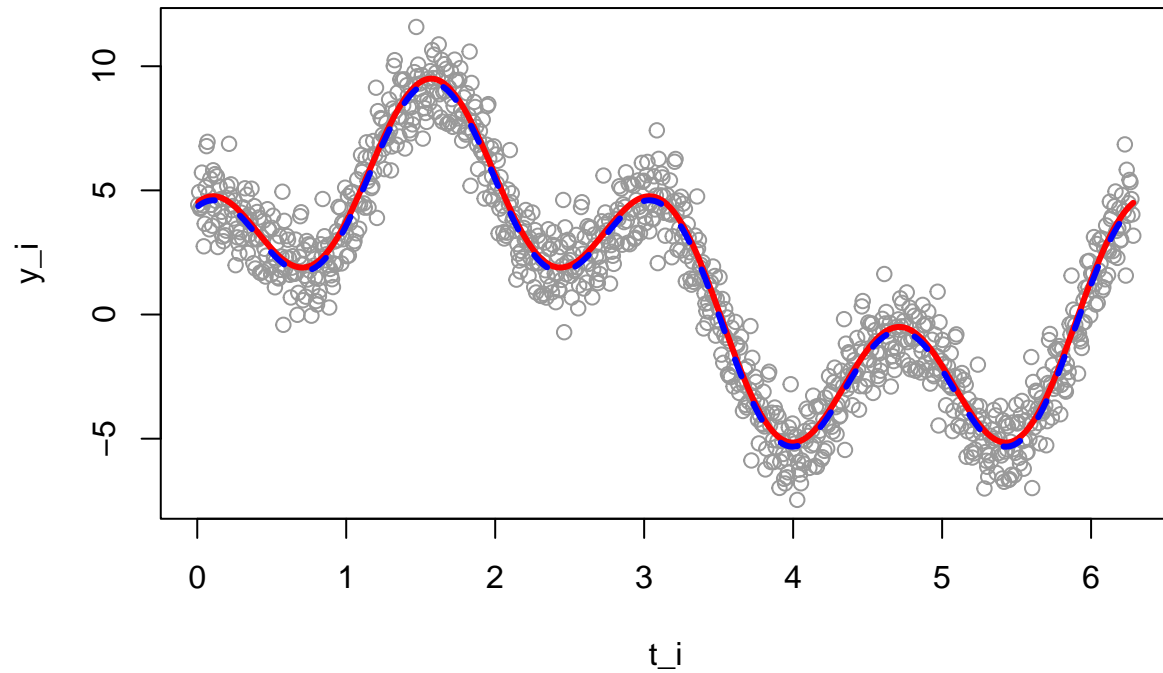
```
##      [,1]
## [1,] 1.5
## [2,] 5.0
## [3,] 3.0
```

(7)

```
XB=X %*% B_
head(XB)
```

```
##      y_i
## [1,] 4.366711
## [2,] 4.395283
## [3,] 4.421961
## [4,] 4.446746
## [5,] 4.469640
## [6,] 4.490648
```

```
plot(t_i,y_i,col='gray60')
lines(t_i,V,col='red',lwd=3)
lines(t_i,XB,col='blue',lty=2,lwd=3)
```



2번

(1)

```
x=seq(from=-1.96, to=1.96, by=0.01)
y=1/sqrt(2*pi)*exp(x^((-1/2)*x^2))
```

```
xx=runif(10000)
xx=xx*2
yy=runif(10000)*1/sqrt(2)
```

```
test=function(xx,yy){
  yy < 1/sqrt(2*pi)*exp(xx^((-1/2)*xx^2))
}
```

```
tst=c()
for(i in 1:10000) tst[i]=test(xx[i],yy[i])

sum(tst)
```

```
## [1] 9670
```

```
sum(tst)/10000
```

```
## [1] 0.967
```

(2)

```
A = rnorm(1000,0,1)
```

```
((-1.96)<=A & A <=(1.96)) %>% sum
```

```
## [1] 970
```

### 3번

-TypeA

```
surv10_prob=0.5^(20)
surv9_prob=c()
for (i in 0:19){
  surv9_prob[i+1]=0.5* 0.95^i * 0.5^(19-i)
}
surv8_prob=c()
for (i in 0:18){
  surv8_prob[i+1]=choose(19-i,1)*0.5*0.95^i * 0.05* 0.5^(18-i)
}
surv10_prob*10+ sum(surv9_prob)*9 + sum(surv8_prob)*8
```

```
## [1] 3.938901
```

- TypeB

```
survprob10=0.5^(20)
survprob9=19*0.5*0.5*0.5^(18)
survprob8=18*0.5*0.5^2*0.5^(17)
survprob10+ survprob9 + survprob8
```

```
## [1] 3.623962e-05
```

### 4번

(1)

```
df=read_csv('https://raw.githubusercontent.com/guebin/2021IR/master/_notebooks/covid19.csv')
```

```
## Rows: 12294 Columns: 5
```

```
## -- Column specification -----
```

```
## Delimiter: ","
```

```
## chr (1): prov
```

```
## dbl (4): year, month, day, cases
```

```
##
```

```
## i Use `spec()` to retrieve the full column specification for this data.
```

```
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
head(df)
```

```
## # A tibble: 6 x 5
```

```
##   year month   day prov  cases
##   <dbl> <dbl> <dbl> <chr> <dbl>
## 1  2020     1    20 서울     0
## 2  2020     1    20 부산     0
## 3  2020     1    20 대구     0
## 4  2020     1    20 인천     1
## 5  2020     1    20 광주     0
## 6  2020     1    20 대전     0
```

```
df %>% group_by(year) %>% summarize(sum(cases))
```

```
## # A tibble: 2 x 2
##   year `sum(cases)`
##   <dbl>     <dbl>
## 1  2020     60726
## 2  2021    396886
```

(2)

```
df %>% filter(year==2021, month==2, day <=15) %>% group_by(prov) %>% summarise(sum(cases))
```

```
## # A tibble: 18 x 2
##   prov `sum(cases)`
##   <chr>     <dbl>
## 1 강원         78
## 2 검역        150
## 3 경기       1708
## 4 경남         95
## 5 경북         84
## 6 광주        166
## 7 대구        183
## 8 대전         49
## 9 부산        278
## 10 서울       2164
## 11 세종        14
## 12 울산        22
## 13 인천       335
## 14 전남        30
## 15 전북        49
## 16 제주        25
## 17 충남       165
## 18 충북        68
```

(3)

```
df %>% filter(year==2021, month==2, day >15) %>% group_by(prov) %>% summarise(sum(cases))
```

```
## # A tibble: 18 x 2
##   prov `sum(cases)`
##   <chr>     <dbl>
## 1 강원         91
## 2 검역         94
## 3 경기       2039
## 4 경남         77
## 5 경북       152
## 6 광주       135
## 7 대구       132
## 8 대전         44
## 9 부산       188
## 10 서울      1916
## 11 세종        16
## 12 울산        54
## 13 인천       283
```

## 14	전남	80
## 15	전북	103
## 16	제주	23
## 17	충남	263
## 18	충북	114