Assignment 2

Wen Yuan Yen

10/23/2018

Part 1: R

6. Compare MCMC outcomes

Test data set

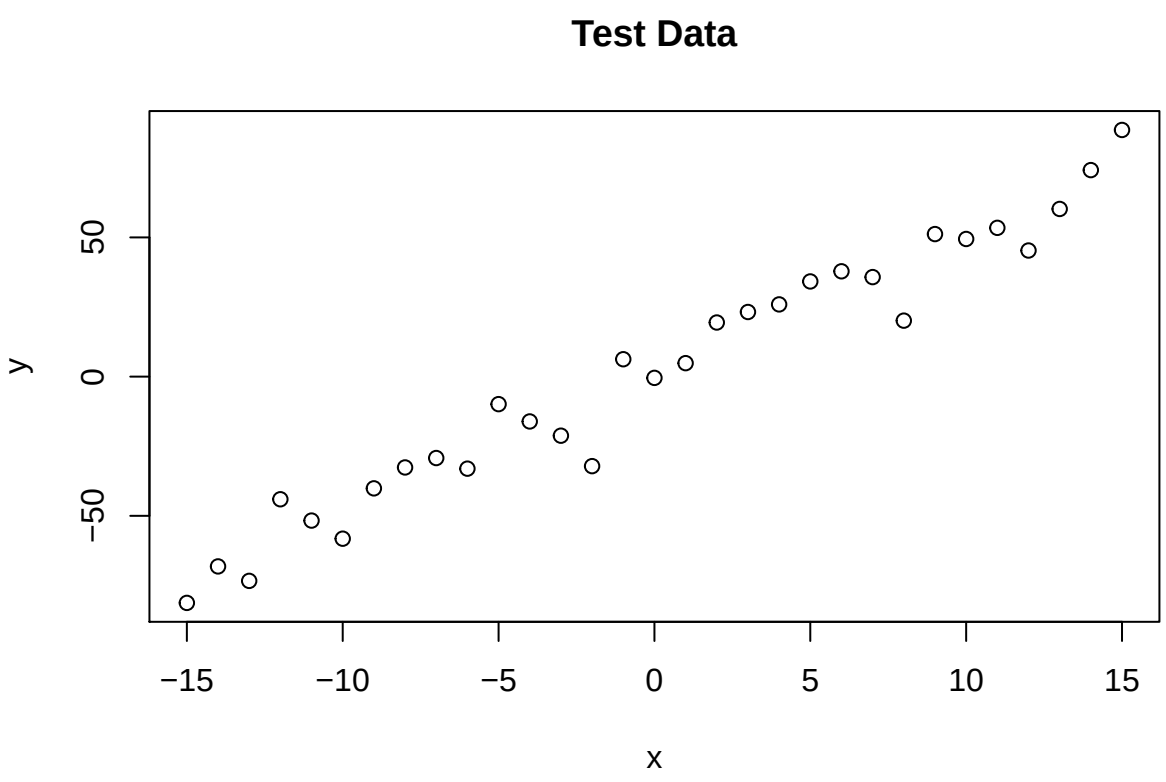
The true model is

$$Y = 5X + \varepsilon, \quad \text{where} \quad \varepsilon \sim N(0, 10^2)$$

```
setwd('~/Desktop')
source('MH algorithm (functions).R')
set.seed(1)

trueA <- 5
trueB <- 0
trueSd <- 10
sampleSize <- 31

x <- (-(sampleSize-1)/2):((sampleSize-1)/2)
y <-  trueA * x + trueB + rnorm(n=sampleSize,mean=0,sd=trueSd)
plot(x,y, main="Test Data")
```



MCMC iterations

When we do burn-in no matter what number of iterations we run, mean level of the chain is quite close to the true parameter. The standard error of the estimator is comparable to that of OLS, which has 0.1913. Hence MCMC is a good alternative to the simple linear regression.

```
m <- seq(1000,100000,length = 10)
compare_outcomes(m, 0.5) # burn-in 50%, in each loop
```

```
##    MCMC.iterations     Mean Std.Error
## 1             1000 5.032702 0.1624983
## 2            12000 4.952581 0.1929331
## 3            23000 4.976845 0.1987602
## 4            34000 4.978114 0.2051908
## 5            45000 4.979188 0.1943114
## 6            56000 4.963502 0.1947384
## 7            67000 4.967163 0.2024883
## 8            78000 4.967612 0.1982249
## 9            89000 4.962829 0.2035333
## 10          100000 4.968679 0.1977119
```

OLS

```
summary(lm(y~x))
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -23.440  -5.055   1.746   6.338  14.376
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.2363     1.7111   0.722    0.476
## x             4.9716     0.1913  25.987   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.527 on 29 degrees of freedom
## Multiple R-squared:  0.9588, Adjusted R-squared:  0.9574
## F-statistic: 675.3 on 1 and 29 DF,  p-value: < 2.2e-16
```