Lecture 02
Solving Problem with R

Oran Kittithreerapronchai[1]

[1]Department of Industrial Engineering, Chulalongkorn University
Bangkok 10330 THAILAND

last updated: August 7, 2023

Outline

source: General references [NC20, TSK16, BG19, Pat14]

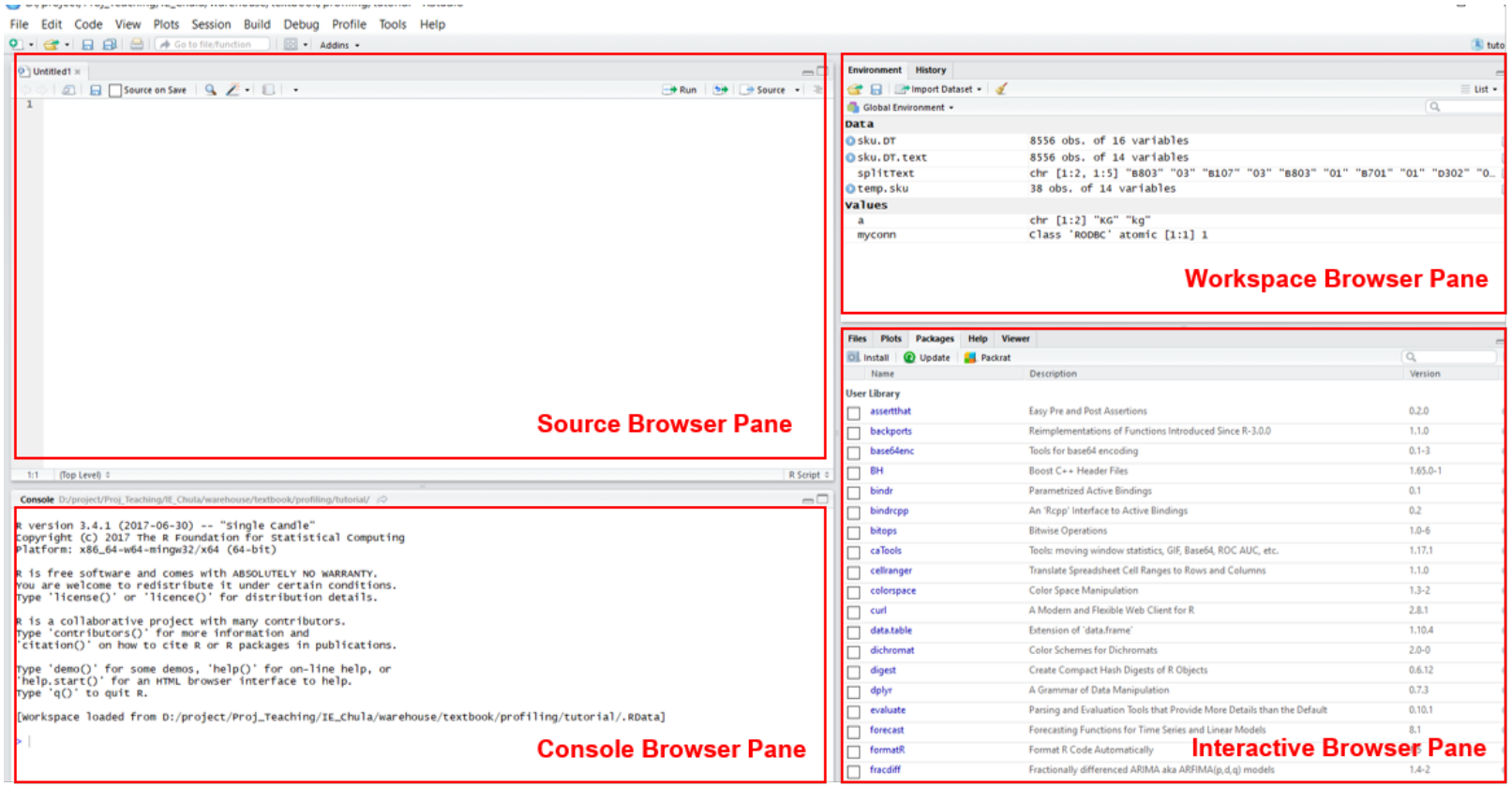
What is R?

- **What:** a free general purposes interpreter program for Statistics and Mathematics
- **Issues:** learning curve, integration, debug
- **Download:** http://mirrors.psu.ac.th/pub/cran/
- **Manual:** http://mirrors.psu.ac.th/pub/cran/manuals.html

R Studio

- **What:** an IDE for R (helps, file explore, workspace)
- **Download:** http://www.rstudio.com/ide/download/desktop

Welcome Screen



RStudio Useful Short Cut

#	add line comment
Ctrl + Shift + C	comment line/ uncomment line
Ctrl + Shift + R	insert code section
Ctrl + Shift + /	reflow comments
Ctrl + Enter	run entire code block
Ctrl + Space	code completion
Ctrl + .	find function that executed
F3 / Shift + F3	Find next/ Find previous; needed Ctrl + F
Ctrl + Shift + Alt + E	find scope of $\{\cdot\}$ or $\cdot$
Ctrl + Shift + A	reflow the code
Ctrl + L	clear console
Ctrl + Tab	go to previous file

Basis Commands in R

- **Algebra:**

```
5 +  3          ; 5 *3              ; 5^3
5 %%  3         ; 5.6 %\% 3
```

- **Vector:**

```
c(1,2) + c(3,4)   ;  2*c(1,2)
c(7,6) + 3        ;  print( c(1,2) * c(3,4) )
```

- **Variable:**

```
varName <- c(1,2)     ## create variable 'varName'
length(varName) ; typeof(varName)    ; str(varName)
varName[1] <- 2
```

- **Sequence:**

```
rm(varName)           ## remove 'varName'
varName <- seq(from=1,to=10,by=2)
sum(varName)
cumsum(varName) + cumprod(varName)
varName <- c(varName,1:sum(varName))
varName > 10
isOdd <- varName %% 2 == 1
varName[isOdd]
```

- **Constant:**

```
month.name   ;  month.abb
LETTERS      ;  letters
pi           ;  Sys.Date()    ; Sys.Time()
```

Basis Commands in R

- **Function:**

```
testFn   <- function(x){ 2*log(x) }
varName2 <- testFn(varName)

emptyFn  <- function(x){}            ## create buffer function
body(emptyFn) <- expression(3*x*sin(3*pi*x)) ## give value
```

- **General:**

```
which.min(varName2)
summary(varName2)
order(varName2,decreasing=F)
isExceed20 <- varName2 > 20
```

- **Matrix:**

```
aMat <- matrix( 1:16, ncol=4)
bMat <- cbind( c(1,5,8,9),c(0,1,-2,0),c(3,5,7,9),c(1,1,1,1))
det(aMat + bMat)
solve(aMat %*% t(bMat) )
```

- **Graphic:**

```
tempSeq <- seq(0,20,0.5)
plot(tempSeq,testFn(tempSeq))
```

- **Statistic:**

```
stdNorm <- rnorm(50)              ## set.seed(17)
hist(stdNorm,breaks=10)
quantile(stdNorm,prob=0.45)
table(floor(10*stdNorm))

shiftUnif <- runif(1000)-0.5
t.test(shiftUnif,rnorm(1000))
shapiro.test(rbinom(5000,1000,0.5))
```

Script File 101

- **Workspace:** all history/object/ instant memory are stored , i.e., *<wd_fileName>.RData*
- **R Object:** A R object can be stored, i.e., *<object>.rds*
- **Rmarkdown:** document with R object and can be convert into html, pdf i.e., *<markdown>.rmd*
- **Script file:** set of commands, i.e., *<script_fileName>.R*
- **Writing Script:** any text editors, i.e., *notepad*, *emac*, jEdit

Example

- **Workspace:**

```
getwd()      ;  setwd("C:/")     ;  R.home()
sink("default.out")              ;  source("scriptFile.R")
rm(ls(all.names=T))              ;  dir()
```

- **Datasets:**

```
saveRDS(iris,"myObj.rds")     ; readRDS("myObj.rds")
save(iris,mtcars,file="multiObj.RData")
save.image("allObj.RData")      ; load("allObj.RData")
```

Example: Engineering Economy

Determine Net Present Value, Payback Period, and Internal Rate of Return of the following cash flow of project if an annual investment interest rate is 8%.

Year	0	1	2	3	4	5	6
Cash flow	-1350	363	551	681	761	821	1467

Recall Given a project cash flow ($\mathbf{cf} = \{cf_0,\ cf_1,\ cf_2, \ldots\ cf_N\}$) and annual interest rate (r), Net Present Value (NPV) is:

$$\text{NPV} = \sum_{i=0}^{N} \left(\frac{cf_0}{(1+r)^i} \right)$$

Life without R

Eng Econ. Trivial

- **NPV:** equivalent cash at time 0; future cash flow is discounted at specific rate
- **Payback period:** period at which cumulative present value = 0
- **IRR:** rate at which NPV = 0

		r=0.08		$r = 0.40$	
i	cf_i	$\frac{cf_i}{(1+r)^i}$	$\sum_{j=0}^{i} \frac{cf_j}{(1+r)^j}$	$\frac{cf_i}{(1+r)^i}$	$\sum_{j=0}^{i} \frac{cf_j}{(1+r)^j}$
0	-1350	-1350.0	-1350.0	-1350.0	-1350.0
1	363	336.1	-1013.9	259.3	-1090.7
2	551	472.4	-541.5	281.1	-809.6
3	681	540.6	-0.9	248.2	-561.4
4	761	559.4	558.5	198.1	-363.3
5	821	558.8	1117.2	152.7	-210.7
6	1467	924.5	2041.7	194.8	-15.8

Engineering Economy as R Script

- **Question:**

```
cashflow <- c(-1350, 363, 551, 681, 761, 821, 1467)
intDef   <- 0.08
```

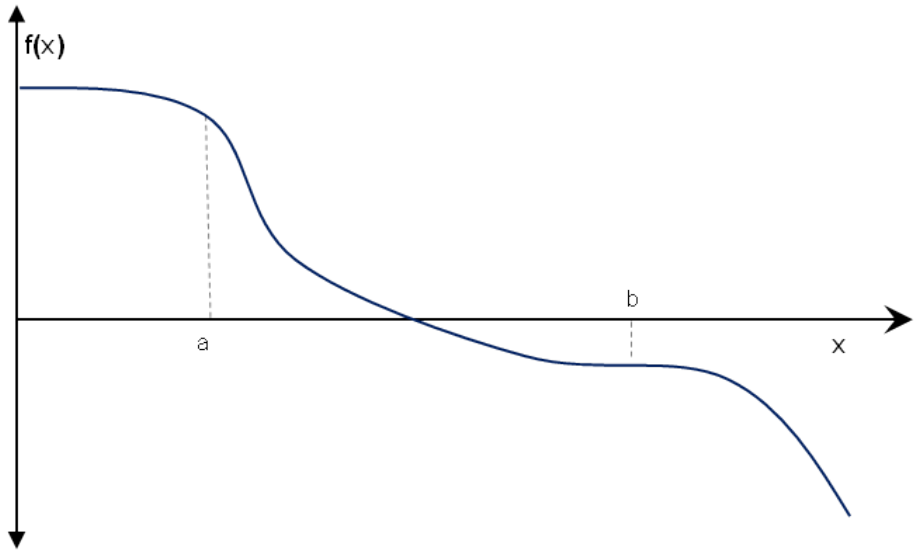
- **NPV:**

```
period <- 0:(length(cashflow) -1)
cashflow.pv <- cashflow /(1+intDef) ^ period
sum(cashflow.pv)
```

- **Payback:**

```
pv.cumsum <- cumsum(findPV(cashflow,intDef))
pv.period <- 0:(length(cashflow)-1)
plot(pv.period,pv.cumsum)
approx(pv.cumsum,pv.period,xout=0.0)
```

Bisection Search: $f(x) = 0$



- compute value at the bounds of interval
- ensure different signs
  - determine the mid point of interval and compute its value
  - ensure difference signs and update the interval
  - repeat until value of mid point ≈ 0

Graphic with R

Using Plot NPV of the project against annual interest rate $r \in [0.01, 0.20]$

- **Calculate:**

```
intRange <- seq(0.01,0.20,0.01)
npvRange <- sapply(intRange,findNPV,cashflow=cashflow)
```

- **Plot:**

```
plot(intRange,npvRange)
abline(h=0,col="red")  ## draw red line at y=0
```

- **Decorate:**

```
plot(intRange,npvRange, cex=0.5, pch=16, type="b"
      ,title="NPV Vs Intestest"   ## specific title
     ,xlim=c(0,0.45),ylim=c(-1200,6000) ## specific scale
     ,xaxt='n',yaxt='n',ylab="Net Present Value",xlab="interest rate") ## label
axis(2,at = seq(-1000,6000,200),col.tick="grey",cex.axis=0.7)
axis(1,at = seq(0,0.4,0.05),col.axis="blue",cex.axis=0.7)
mtext(4,text = paste("created by OKC on ",Sys.Date()),cex=0.6)
text(0.2,4200, paste( c("cash flow = ",cashflow),collapse=" "),cex=0.8)
```

Useful R commands

- **I/O:**

```
write.table(cbind(intRange,npvRange),file="output.csv",row.names=T,col.names=T,sep=" ")
npvData <- read.csv(file = "output.csv",multi.line=T)
str(npvData)
```

- **Calculus:**

```
expr <- expression( x^2 + 3* sin(x) + log(x) )  ;  D(expr,"x")

exprFn <- function(x){}
body(exprFn) <- deriv3(expr,"x")                ;  exprFn(3)

integrate( function(x){x^2},0,1)
```

What is bond

- **Bond:** long-term note issued by business or government in exchange of coupon and principle when maturity
- **Type:** zero coupon bonds, floating rate debt, junk bonds, semiannually
- **Provision:** call provision (cooperate buys), redeemable (investor sells), sinking fund (shrinking value)
- **Rating:**

	Moody's	S&P	Fitch
High grade	Aaa	AAA	AAA
Medium grade	Baa2	BBB	BBB
Non-investment	Ba1	BB+	BB+
Speculative	B2	B	B
Junk	C	C	CCC

- **Value:** regular NPV (**yield to maturity** and **yield to call**)

In Class Exercise

Bond Valuation

A 10-year, 12 % semiannual coupon bond, with a par value of $1000 may be called in 4 years at a call price of $1060. The bond sells for $1100 (Assume that the bond has just been issued.)

- What is the bond's yield to maturity?
- What is the bond's yield to call?

source: Brigham, E *et al.*. 2000 Problem 8-9

Cooperate Bond Spread

Consider a 12-Years 7.980000% of annual coupon bond of BANPU plc issued on 01-Apr-11 with a par value of THB1000. Answer the following questions

- If such bond trades at THB1020 on 01-Apr-11, should we buy it?
- If such bond trades at THB980 on 02-Apr-13, should we buy it?
- What is the spread of this bond comparing to Thai bonds (Yield Curve: `thYeild-20130401.csv`)?

Useful URL Link

Data

- **Derivative:** http://www.tfex.co.th/tfex/historicalMarketReport.html
- **Bonds:** http://www.thaibma.or.th/PriceYield.html
- **Stock:** http://161.200.66.44:8080/loginbypass.jsp
- **Others** http://www.ie.eng.chula.ac.th/õran/classes/FinOpt_resources/data/

R

- **Wiki:** http://en.wikibooks.org/wiki/R_Programming
- **Tutorial:** http://www.statmethods.net/graphs/boxplot.html
  http://www.cyclismo.org/tutorial/R/index.html
- **StackOverFlow:** http://stackoverflow.com/questions/tagged/r

Reference

[BG19] Brad Boehmke and Brandon M Greenwell.
Hands-on machine learning with R.
CRC press, 2019.

[NC20] Fred Nwanganga and Mike Chapple.
Practical machine learning in R.
John Wiley & Sons, 2020.

[Pat14] Manas A Pathak.
Beginning data science with R.
Springer, 2014.

[TSK16] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar.
Introduction to data mining.
Pearson Education India, 2016.