



R Introduction

胡翠華

2005/9/30

- 下載、啟動與離開
- 環境介面
- 指令語法
- 資料型態、存放方式
- 流程控制
- 圖形
- 讀寫檔案
- 儲存結果
- 自訂函數

下載：<http://www.r-project.org/>

The R Project for Statistical Computing

About R
[What is R?](#)
[Contributors](#)
[Screenshots](#)
[What's new?](#)

Download
[CRAN](#)

R Project
[Foundation](#)
[Members & Donors](#)
[Mailing Lists](#)
[Bug Tracking](#)
[Developer Page](#)
[Search](#)

Documentation
[Manuals](#)
[FAQs](#)
[Newsletter](#)
[Books](#)
[Other](#)

http://cran.csie.ntu.edu.tw/
Windows (95 and later) :
/base/rw2011.exe

中文手冊

Clustering 4 groups

Factor 1 [41%]

Factor 3 [19%]

ing Started:

- R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To download R, please choose your preferred [CRAN mirror](#).

http://www.r-project.org/misc/acplust.R

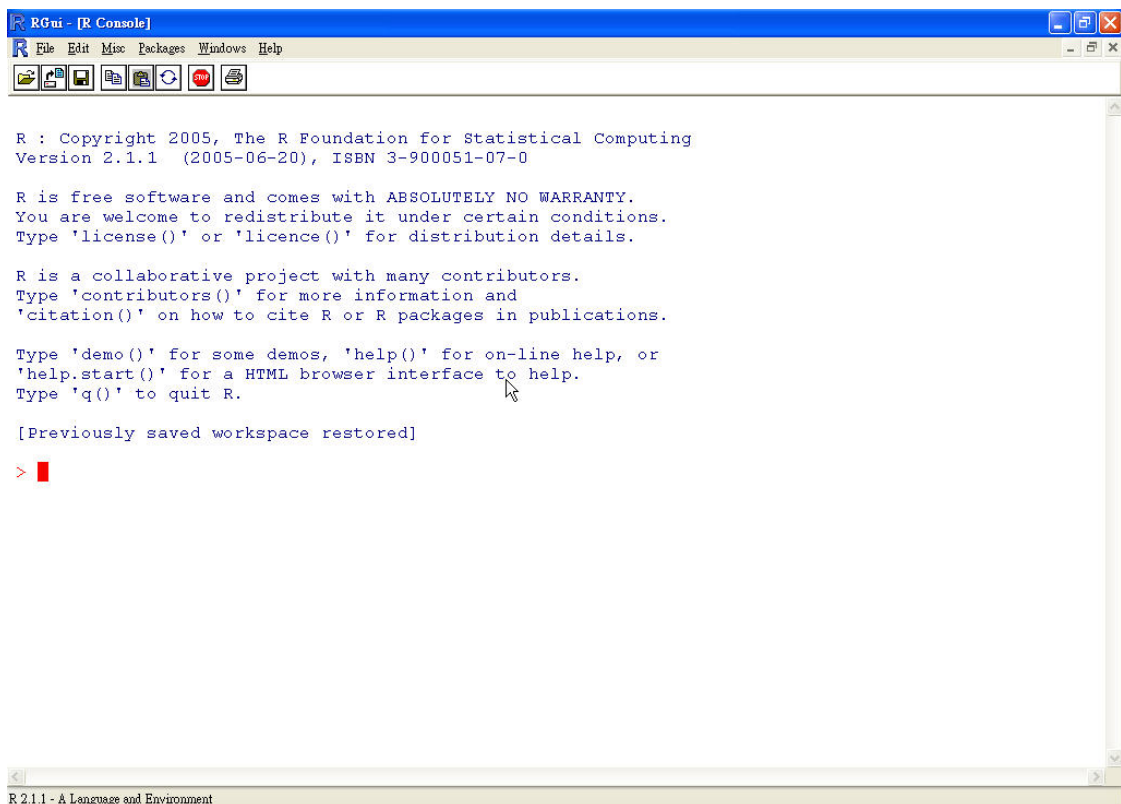
網際網路

■ 安裝

- 安裝路徑 C:\Program Files\R\rw2011
- 25 packages in the ... \R\rw2011\library*
- 內設工作環境:
... \R\rw2011\.RData

■ 啟動

- ... \R\rw2011\bin\Rgui.exe
windows 介面(右圖)
- ... \R\rw2011\bin\
 - R.exe
 - Rterm.exe如同 UNIX 環境介面



```
RGui - [R Console]
File Edit Misc Packages Windows Help
[Icons]

R : Copyright 2005, The R Foundation for Statistical Computing
Version 2.1.1 (2005-06-20), ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for a HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> █
```

■ 啟動

- 7 packages loaded

> search()

[1] ".GlobalEnv" "package:methods" "package:stats"

[4] "package:graphics" "package:grDevices" "package:utils"

[7] "package:datasets" "Autoloads" "package:base"

■ 離開

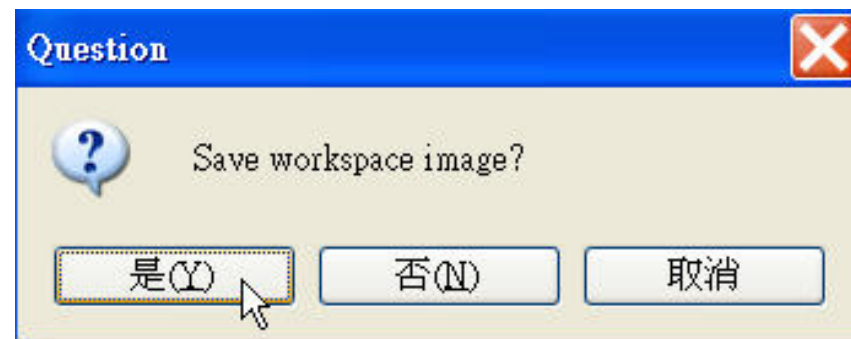
- File → Exit

- > q()

- command視窗右上方 

■ save workspace

- 內設.../R/rw2011/.RData



環境介面：**command**視窗

■ 簡單的數值運算

```
> 1 + 3
```

```
[1] 4
```

```
> (10.4 - 5.2) * 30 / 21
```

```
[1] 7.428571
```

```
> sin(pi/2) ; log(10)
```

```
[1] 1
```

```
[1] 2.302585
```

■ 複雜的運算、程式編輯與執行...

環境介面：File

Source R code...

... 讀進已經建立的檔案

New script

... 編輯程式、文件

Open script...

... 開啟已建立的檔案

Display file(s)...

... 顯示已建立的檔案內容

Load Workspace...

... 引入其他工作環境

Save Workspace...

... 儲存主要工作環境

Load History...

... 顯示歷史記錄

Save History...

... 儲存歷史記錄

Change dir...

... 變更工作環境路徑

Print...

... 列印

Save to File...

... 儲存成檔案

Exit

... 離開R

環境介面：Edit

<u>C</u> opy	Ctrl+C
<u>P</u> aste	Ctrl+V
Paste commands only	
<u>C</u> opy and <u>P</u> aste	Ctrl+X
<u>S</u> elect all	
<u>C</u> lear console	Ctrl+L
<hr/>	
<u>D</u> ata editor...	
<hr/>	
<u>G</u> UI preferences...	

... 複製

... 貼上

... ?

... 複製與貼上

... 選擇command視窗上全部的畫面

... 清除**command**視窗上全部的畫面

...編輯已存在的R資料檔

...設定介面(**command**視窗大小、字型與大小...)

暫時

永久 — 儲存 .../R/rw2011/etc/Rconsole

環境介面：Misc

Stop current computation	ESC
✓ Buffered output	Ctrl+W
List objects	
Remove all objects	
List search path	

... 中斷正在執行的工作

... 不勾選表示執行過程立即顯示結果

... 顯示主要工作環境下的所有檔案(1)

... 移除主要工作環境下的所有檔案 (2)

... 顯示目前已引進的 **package** 或物件或其他工作環境(3)

Note:

(1) **ls(), objects()**

(2) **rm(list=ls(all=T))**
remove(list=ls(all=T))

(3) **search()**

環境介面：Package

Load package...

Set CRAN mirror...

Select repositories...

Install package(s)...

Update packages...

Install package(s) from local zip files...

- ... 引進package (已下載的package) (1)
- ... 設定下載package的網站
- ... 選擇下載package的地方(CRAN, ...)
- ... 從網路上安裝package
- ...更新packages(電腦中已經有的package)
- ... 從CRAN網站上下載packages的壓縮檔
來執行安裝packages

Note:

- (1) **library(package name),**
require(package name)

環境介面：其他

■ Windows

- 多個視窗的排列方式

■ Help

- 說明文件
- 連結R homepage
- 查詢指令 (Note:)

...

Note:

```
(1) ?plot, help("plot")    #知道指令名稱  
    help.search("plot") #不知道指令名稱  
    apropos("plot") #函數名稱中有'plot'字串的所有函數  
    find("plot") # plot在哪個package  
    getAnywhere("plot") # plot函數出現在哪些package
```

```
(2) search from http://search.r-project.org  
Help -> search.r-project.org...  
RSiteSearch
```

package: 需已下載

工作環境

■ 目前的環境

主要工作環境

```
> search()
```

```
[1] ".GlobalEnv"      "package:methods" "package:stats"  
[4] "package:graphics" "package:grDevices" "package:utils"  
[7] "package:datasets" "Autoloads"      "package:base"
```

```
> searchpaths()
```

```
[1] ".GlobalEnv"  
[2] "C:/PROGRA~1/R/rw2011/library/methods"  
[3] "C:/PROGRA~1/R/rw2011/library/stats"  
[4] "C:/PROGRA~1/R/rw2011/library/graphics"  
[5] "C:/PROGRA~1/R/rw2011/library/grDevices"  
[6] "C:/PROGRA~1/R/rw2011/library/utils"  
[7] "C:/PROGRA~1/R/rw2011/library/datasets"  
[8] "Autoloads"  
[9] "C:/PROGRA~1/R/rw2011/library/base"
```

工作環境

■ 主要工作環境

□ 內設的路徑在 C:/Program Files/R/rw2011

■ 查詢指令 `getwd()`

□ `command` 視窗下的物件儲存在

`.../R/rw2011/.RData`

■ 查詢物件: `ls`, `objects`

■ 刪除物件: `rm`, `remove`

工作環境

■ 其他工作環境

□ Load Workspace：工作環境檔*.RData需已存在

■ `load("d:/test.RData")`

■ `attach("d:/test.RData", pos=2)` 加入

`detach(pos=2)` 移除

□ Save Workspace

■ `save.image(".RData", compress=F)`

`save.image("d:/test.RData ", compress=F)`

■ `save(list=c("data1", "data2"), file="d:/test.RData", compress=F)`

基本指令語法

- 名稱<-物件 或 名稱=物件($\geq R1.4.0$) 或 物件->名稱

物件=名稱 \Rightarrow 錯誤

□ 名稱：

- 英文字母(字串)、英文子母(字串)與數字、"."、"_"的組合

- 英文大小寫字母有區別

```
> y<- 4
```

```
> Y<- 40
```

```
> x3_4z<- 1.32
```

- "_"不可在第一個位置

- "."在第一個位置

```
> .x<- 3 ; y<-5
```

```
> ls()
```

```
[1] "y"
```

```
> ls(all.names=T)
```

```
[1] ".x" "y"
```

基本指令語法

□ 物件

■ 數值、字元...

```
> x1<- 3.0
```

```
> x2<- "a"
```

■ 向量、矩陣...

```
> y1<- c(1, 2.3, 0.23)
```

```
> y2<- matrix(c(1,2,3,4,5,6), nrow=3, ncol=2)
```

■ 執行R function的結果

```
> x.lm <- lm(Y~X1+X2, data=mydata)
```

■ 自訂的函數(function)

基本指令語法

■ R指令

□ function

■ 格式：**function(參數){ ... }**

■ 查看內容(...): **> fun_name**

□ 查詢用法

例子:

>?lm

lm	package:stats	R Documentation
Description:		
Usage: lm(formula, data, subset, weights, na.action, method = "qr", model = TRUE, x = FALSE, y = FALSE, qr = TRUE, singular.ok = TRUE, contrasts = NULL, offset, ...)		
Arguments:		
Value:		
...		
See Also:		
Examples:		

基本指令語法

□ 查詢參數用法

例子:

```
> args("lm")
```

```
function (formula, data, subset, weights, na.action, method  
= "qr", model = TRUE, x = FALSE, y = FALSE, qr =  
TRUE, singular.ok = TRUE, contrasts = NULL,  
offset, ...)
```

參數—

- 必要參數、非必要參數
- 順序不可對調：參數名稱可不給定 ex. `lm(y~x, mydata)`
- 順序可對調：參數名稱給定 ex. `lm(data=mydata, formula=y~x)`

基本指令語法

- ~ (formula: $Y \sim X$ 表示 $Y = aX + b$)

```
> y.lm <- lm(y~x, data=mydata)
```

- mean, median, sum, var, sd, quantile, summary

```
> summary(c(1,4,2,3,2,2,5))
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.000	2.000	2.000	2.714	3.500	5.000

```
> summary(y.lm)
```

- l(大寫i)

- ex1. $l(X < 2)$

- ex2. $lm(y \sim x + l(x^2), data)$

資料型態

- 數值：numeric (is.numeric, as.numeric)
 - 整數：integer (is.integer, as.integer) $-2 \cdot 10^9 \sim 2 \cdot 10^9$
 - 實數：real (is.real, as.real)
 - x1<- 3
 - x2<- 3.0
- 字元：character (is.character, as.character)
 - x3<- "3" # x3<- '3'
- 複數：complex (is.complex, as.complex)
 - x4<- 3+2i

資料型態

- 邏輯：logical (is.logical, as.logical)
x5<- TRUE # x<- T
y6<- FALSE # y<- F
- 類別：factor(x, ordered=F) (is.factor, as.factor)
x7<- factor(c("a", "g", "e"), ordered=T)
- 日期型態：as.Date(x, format="%Y-%m-%d")
 - 例如2003年3月25日
as.Date("2003/3/25")
as.Date("3/25/2003", "%m/%d/%Y")
as.Date("03/25/03", "%m/%d/%y")
R中顯示為"2003-03-25"
 - months, weekdays

資料型態

■ 附註

- 資料型態通常不需要先宣告，例外
 - 類別 (factor)、日期 (as.Date)
 - R與Fortran和C連結時，資料的傳遞必須先指定資料的型態，否則可能造成執行結果錯誤或中斷程式
- 遺漏值(missing)： NA (is.na), NaN (is.nan, is.na)
 - > x<- NA ; is.na(x) ; !is.na(x)
 - [1] TRUE
 - [1] FALSE
 - > sqrt(-2)
 - [1] NaN
- is.* 是否為...
as.* (強迫)使為...

資料型態

■ 資料型態之間的轉換

例1.

```
> x1<- 3
```

```
> y1<- as.character(x) # y<-"3"
```

```
> x2<-"test" #x為字元型態
```

```
> y2<- as.factor(x2) #將字元型態x轉成類別型態
```


資料型態

■ 資料型態之間的轉換

例2.

```
> x<- as.Date("2003/3/23")
```

```
> as.character(x)
```

```
[1] "2003-03-23"
```

```
> as.numeric( substring( as.character(x), 1, 4 ) ) # 年份
```

```
[1] 2003
```

```
> as.numeric( substring( as.character(x), 9, 10 ) ) # 日
```

```
[1] 23
```

```
> months(x)
```

```
[1] "三月"
```

```
> as.numeric( substring( as.character(x), 6, 7 ) ) # 月
```

```
[1] 3
```

資料存放方式 (向量)

- 向量vector：資料型態需一致

```
> x1<- c(1,4,2,1,1,2)
```

```
> names(x1)<-c("A","B","C","D","E","F") # names(x1)<- LETTERS[1:6]
```

```
> x1
```

```
A B C D E F
```

各元素的名稱

```
1 4 2 1 1 2
```

```
> length(x1)
```

```
[1] 6
```

```
> x1[1] # 同x1[c(1)], x1["A"]
```

```
A
```

```
1
```

```
> x1<- unname(x1) ; x1 # # names(x1) <- NULL
```

```
[1] 1 4 2 1 1 2
```

```
> x1[c(2,3,4,5,6)] #同x1[-1] 或 x1[2:length(x1)]
```

```
[1] 4 2 1 1 2
```

資料存放方式 (向量)

```
> x1[x1==1]
[1] 1 1 1
> which(x1==1)
[1] 1 4 5
```

```
> unique(x1)
[1] 1 4 2
> table(x1) # 向量
```

1 2 4

3 2 1

x1中的數值

出現的次數

資料存放方式 (向量)

```
> x1.new<- paste("a", x1, sep="") # 字串與數值的結合
```

```
> x1.new
```

```
[1] "a1" "a4" "a2" "a1" "a1" "a2"
```

```
>paste(x1.new, collapse=":", sep="")
```

```
[1] "a1:a4:a2:a1:a1:a2"
```

```
> c(x1, c(10,11)) # 向量的合併
```

```
[1] 1 4 2 1 1 2 10 11
```

資料存放方式 (矩陣)

■ 矩陣 `matrix`, `array` : 資料型態需一致

```
> x2<- matrix(c(1,2,3,4,5,6,7,8),nrow=2,ncol=4)
```

```
> x2
```

```
      [,1] [,2] [,3] [,4]  
[1,]  1   3   5   7  
[2,]  2   4   6   8
```

```
> x3 <- matrix(c(1,2,3,4,5,6,7,8),nrow=2,ncol=4, byrow=T)
```

```
> x3
```

```
      [,1] [,2] [,3] [,4]  
[1,]  1   2   3   4  
[2,]  5   6   7   8
```

資料存放方式 (矩陣)

取出其中的元素:

```
> x2[2, 3] #第2列, 第3行的值
```

```
[1] 6
```

```
> x2[2, c(2, 4)] #第2列, 第2, 4行的值
```

```
[1] 4 8
```

```
> x2[2, ] #第2列所有元素的值, 同x2[2, 1:ncol(x2)] , nrow, ncol: 矩陣列數, 行數
```

```
[1] 2 4 6 8
```

```
> dimnames(x2)<- list(paste("r",1:nrow(x2), sep=""), paste("c",1:ncol(x2),  
  sep=""))
```

```
> x2
```

```
  c1 c2 c3 c4
```

```
r1 1 3 5 7
```

```
r2 2 4 6 8
```

```
> x2<- unname(x2) # 除去名稱
```

資料存放方式 (矩陣)

row, column 合併 X2, X3 ($\dim(X2)=\dim(X3)= 2 \times 4$):

```
> cbind(x2,x3)
```

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]
[1,]	1	3	5	7	1	2	3	4
[2,]	2	4	6	8	5	6	7	8

```
> rbind(x2,x3)
```

	[,1]	[,2]	[,3]	[,4]
[1,]	1	3	5	7
[2,]	2	4	6	8
[3,]	1	2	3	4
[4,]	5	6	7	8

資料存放方式 (data frame)

- data frame : 同一行資料型態不需要一致

```
> y<- data.frame(X1=c(1,4), X2=letters[1:2],  
                X3=as.Date(c("2004-1-1","2005-3-4"))) )
```

```
> y
```

	X1	X2	X3
1	1	a	2004-01-01
2	4	b	2005-03-04

```
> y$X3      # y[,3], y[, "X3"]
```

```
[1] "2004-01-01" "2005-03-04"
```

```
> y$X3[2]   # y[2,3]
```

```
[1] "2005-03-04"
```

```
> y[,c(1,3)] # y[1:nrow(y), c(1,3)]
```

	X1	X3
1	1	2004-01-01
2	4	2005-03-04

資料存放方式 (data frame)

data.frame 合併

```
> z1<- data.frame(Z1=c("123","456"), Z2=5:6)
```

```
> cbind(z=y, z1) # data.frame(y, z1)
```

	X1	X2	X3	Z1	Z2
1	1	a	2004-01-01	123	5
2	4	b	2005-03-04	456	6

```
> z2<- data.frame(X1=3:4, X2=letters[3:4], X3=as.Date(c("2005-5-1", "2005-5-2")))
```

```
> w<-rbind(y,z2) # y, z1 行的屬性相同
```

	X1	X2	X3
1	1	a	2004-01-01
2	4	b	2005-03-04
11	3	c	2005-05-01
21	4	d	2005-05-02

資料存放方式 (data frame)

```
> dimnames(y)
```

```
> unname(y)
```

```
c("1", "2") c("a", "b") c("2004-01-01", "2005-03-04")
```

```
1      1      a      2004-01-01
```

```
2      2      b      2005-03-04
```

```
> w<-w[order(w$X1, w$X2, w$X3),]
```

```
      X1  X2      X3
```

```
1     1  a  2004-01-01
```

```
11    3  c  2005-05-01
```

```
2     4  b  2005-03-04
```

```
21    4  d  2005-05-02
```

```
> w[(w$X1==4) | (w$X1==1),]
```

```
□ merge
```

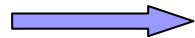
資料存放方式 (list)

- list : 資料型態或存放方式不需要一致

```
> z<- list(A=c(1,2,3), B=matrix(c("a", "b", "c", "d"),nrow=2,ncol=2) )
```

```
> z
```

```
$A  
[1] 1 2 3
```



```
z[[1]] 或 z[["A"]] 或 z$A
```

```
$B  
  [,1] [,2]  
[1,] "a" "c"  
[2,] "b" "d"
```



```
z[[2]] 或 z[["B"]] 或 z$B
```

```
> names(z)
```

```
[1] "A" "B"
```

```
> length(z)
```

```
[1] 2
```

資料存放方式 (list)

```
> unname(z)
```

```
[[1]]
```

```
[1] 1 2 3
```

```
[[2]]
```

```
  [,1] [,2]
```

```
[1,] "a" "c"
```

```
[2,] "b" "d"
```

- output常用的儲存方式
- dimnames: matrix, array, data frame

資料存放方式 (list)

list 合併

```
> w<- list(c=1:5, d=data.frame(V1=1:2, V2=c("e","f")))
```

```
> c(z, w)
```

```
$A
```

```
[1] 1 2 3
```

```
$B
```

```
      [,1] [,2]
```

```
[1,] "a" "c"
```

```
[2,] "b" "d"
```

```
$c
```

```
[1] 1 2 3 4 5
```

```
$d
```

```
  V1 V2
```

```
1  1  e
```

```
2  2  f
```

流程控制

■ 條件語句

```
if(條件){  
    ...  
    ...  
}  
else if(條件){  
    ...  
    ...  
}  
else{  
    ...  
    ...  
}
```

```
if(條件){  
    ...  
    ...  
}
```

```
ifelse(條件, v1, v2)
```

Note: 條件成立為v1
否則為v2

流程控制

■ 迴圈

```
for(name in e1){  
  ...  
}
```

e1: 向量 1:10 , seq(1,10,2)
c("a","b","c")

```
repeat{  
  ...  
  if(條件) break  
}
```

```
while(條件){  
  ...  
}
```

■ break

- 適用於結束任何的迴圈
- 結束repeat唯一方式

■ next

流程控制

■ 避免迴圈

□ 方式

- apply
- sapply
- lapply
- call C, Fortran

□ 例子: 求矩陣X中每一行的平均值

```
> X<- matrix(runif(1000), nrow=10, ncol=100)
> X.m1<- apply(X, 2, FUN="mean", na.rm=T)
> X.m2<- numeric()
> for(i in 1:ncol(X)){
    X.m2[i]<- mean(X[,i])
}
```


圖形

■ 開啟圖形視窗

- `win.graph(width=7, height=7) #win.graph(w=7, h=7)`
- `windows`

■ high-level

- `plot, boxplot, hist, matplot...`

```
> x<- rnorm(100)
```

```
> plot(x)
```

```
> y<- rnorm(100, mean=3)
```

```
> matplot(1:length(x), cbind(x, y)) #x, y長度一樣
```

圖形

■ low-level

已經有 high-level 的圖

- points, lines, abline, points, text, segments...
- title, mtext
- axis
- legend
- expression, plotmath

圖形

■ par

- mfrow, mfcol

- type, lty, lwd, pch

- col (col.lab, col.axis, col.main, col.sub)

- 數字: 1, 2, 3, ...

- > palette()

- [1] "black" "red" "green3" "blue" "cyan" "magenta" "yellow" "gray" # col=1~8

- > palette(c("red", "blue")) # col=1, 2

- > palette("default")

- 字元: "red", "blue", ...

- cex (cex.lab, cex.axis, cex.main, cex.sub)

- las, xaxt, yaxt

- xlim, ylim

...

- 關閉圖形視窗 dev.off(), graphics.off()

圖形

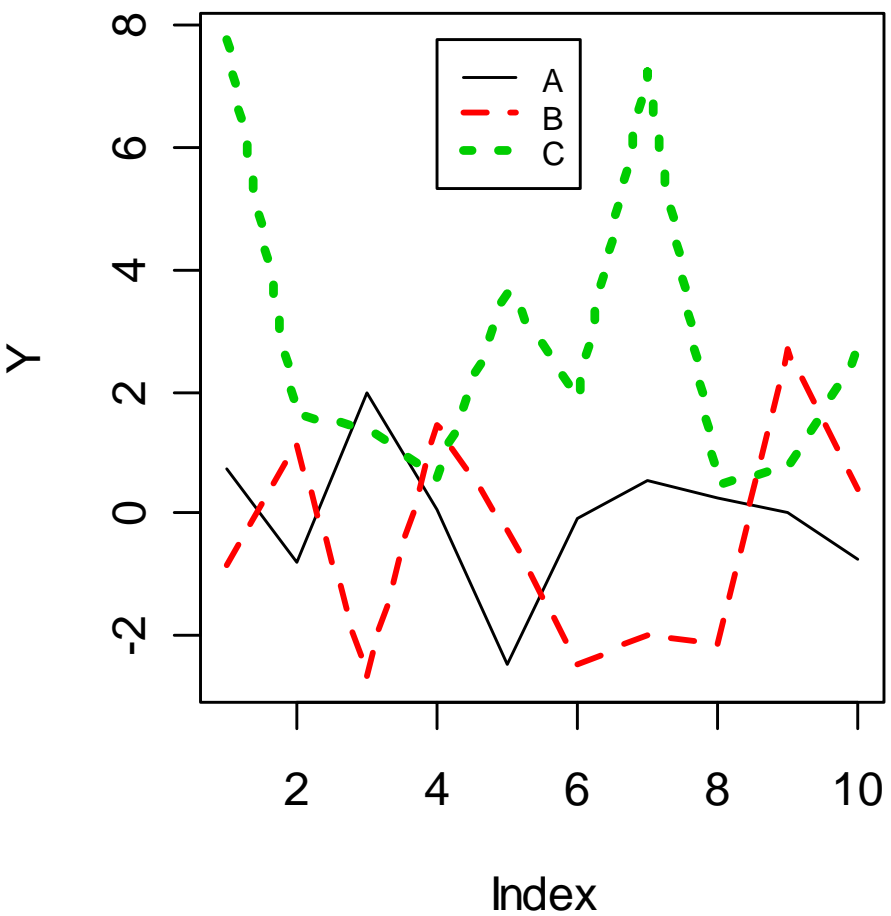
■ 例子1.

```
y<- matrix(c(rnorm(10), runif(10, -3, 3), rchisq(10, 3)), ncol=3)
x<- 1:nrow(y)
win.graph(w=7, h=4)
par(mfrow=c(1,2), las=0)
plot(range(x), range(y), xlab="Index", ylab="Y", type="n")
mtext(expression(graph ^{1}), line=0.5, cex=1)
for(i in 1:ncol(y)){ lines(x, y[,i], col=i, lty=i, lwd=i) }
legend(4, max(y), c("A", "B", "C"), lty=c(1,2,3), lwd=1:3, col=1:3, cex=0.7)

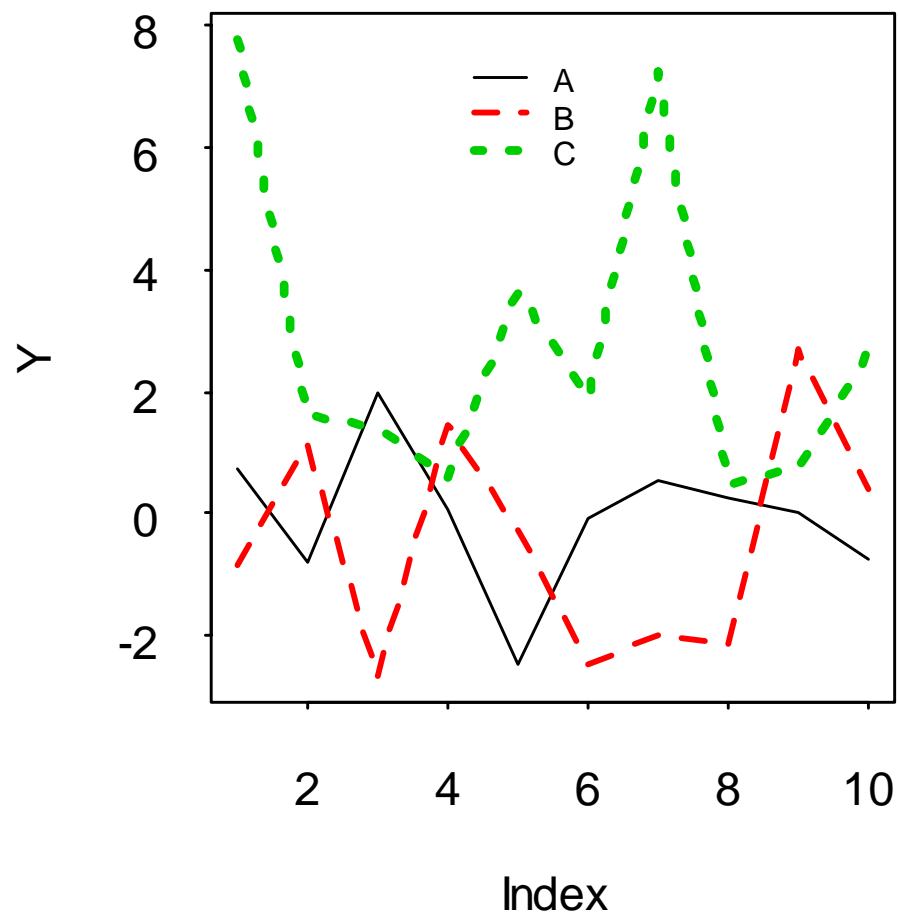
par(las=1, tck=-0.01)
matplot(x, y, xlab="Index", ylab="Y", type="l", col=1:3, lwd=1:3)
mtext(expression(graph [2]), line=0.5, cex=1)
legend(4, max(y), c("A", "B", "C"), lty=1:3, lwd=1:3, col=1:3, bty="n",cex=0.7)
```

圖形

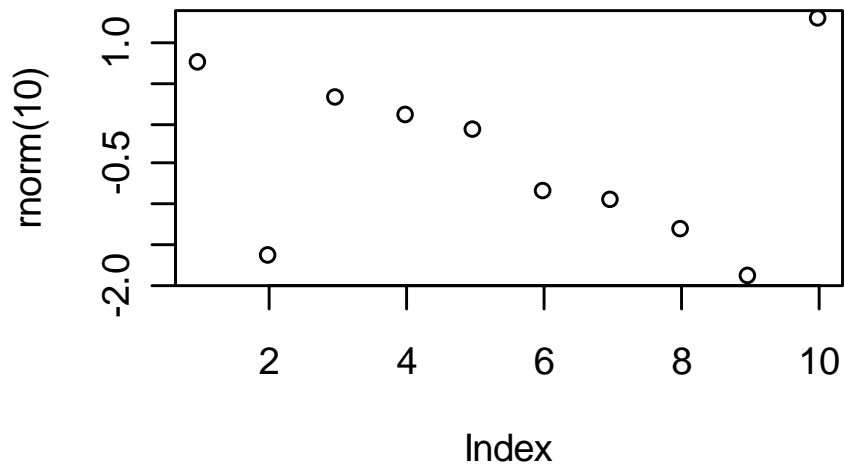
graph¹



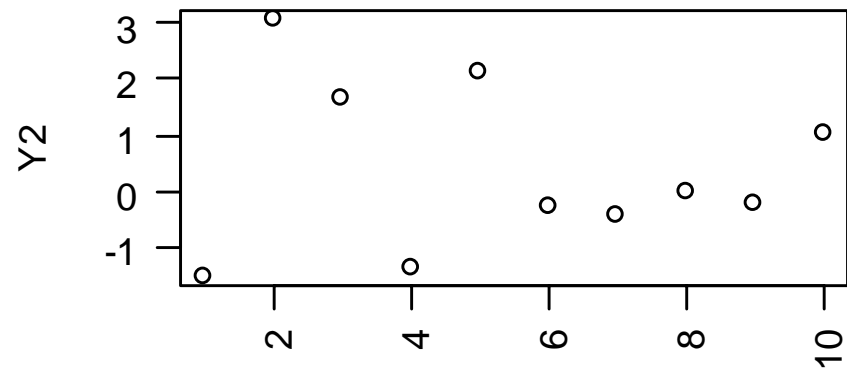
graph₂



A

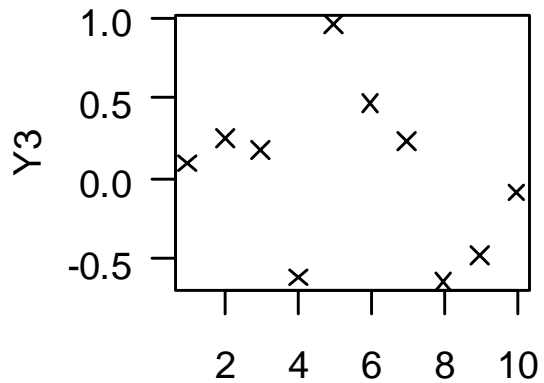


B

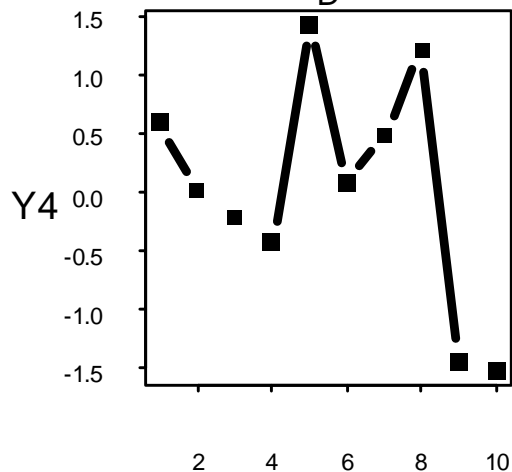


****sub-title****

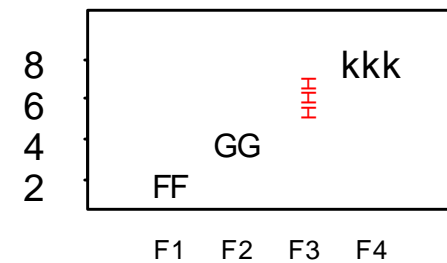
C



D



F



```
par(mfrow=c(2,2), las=0)
par(fig=c(0,0.5,0.5,1)) ; plot(rnorm(10), main="A")
```

```
par(fig=c(0.5,1,0.5,1), new=T, las=2)
plot(rnorm(10), sub="**sub-title**", xlab="", ylab="Y2",main="B", cex.main=0.8,col.main=2)
```

```
par(fig=c(0,1/3,0,0.5), new=T, las=1)
plot(rnorm(10), pch=4, cex=1, xlab="", ylab="Y3") ; mtext("C", line=1, cex=1)
```

```
par(fig=c(1/3,2/3,0,0.5), new=T, mar=c(5,4,2,2), tck=-0.02)
plot(rnorm(10), type="b", pch=15, lty=1, lwd=3, cex.axis=0.6, xlab="", ylab="")
mtext("D", line=0.1, cex=0.7) ; mtext("Y4", side=2, line=2, cex=0.8)
```

```
par(fig=c(2/3,1,0,0.5), new=T, mar=c(5,4,6,2))
plot(c(1,2,3,4), c(2,4,6,8),type="n", xlab="", xaxt="n", yaxt="n", ylab="",xlim=c(0,5),
      ylim=c(1,10))
text(c(1,2),c(2,4),c("FF","GG"), cex=0.8)
par(srt=90) ; text(3,6,"HHH", cex=0.5, col=2) ; par(srt=0) ; text(4,8,"kkk", cex=1, col=1)
axis(side=2, at=c(2,4,6,8),c("2","4","6","8"))
mtext(paste("F",1:4,sep=""), at=c(1,2,3,4), side=1, line=0.2, cex=0.6)
title("F",line=1)
```

讀寫檔案

■ 讀取外部資料檔

□ *read.table*

```
> x<- read.table("d:/data1.txt", header=F, sep=" ",  
  colClasses=c("numeric","character","Date"))
```

不給定colClasses則B, C為factor

```
> x
```

	V1	V2	V3
1	1	a	2003-12-11
2	2	b	2003-12-23
3	3	c	2003-12-25

外部檔：d:/data1.csv

```
1 a 2003/12/11  
2 b 2003/12/23  
3 c 2003/12/25
```


讀寫檔案

□ ***read.fwf*** : 固定寬度

```
> y<- read.fwf("d:/data1.txt", widths=c(2,2,10), sep="",  
  colClasses=c("numeric","character","Date"))
```

```
> y
```

	V1	V2	V3
1	1	a	2003-12-11
2	2	b	2003-12-23
3	3	c	2003-12-25

讀寫檔案

□ *scan*

```
> z<- scan("d:/data1.txt", what=list(numeric(), character(), character()))
```

Read 3 records

```
[[1]]
```

```
[1] 1 2 3
```

```
[[2]]
```

```
[1] "a" "b" "c"
```

```
[[3]]
```

```
[1] "2003/12/11" "2003/12/23" "2003/12/25"
```

```
> z[[3]]<-as.Date(z[[3]], format="%Y/%m/%d")
```

讀寫檔案

```
> w<- scan("d:/data2.txt", what=numeric())
Read 5 items
> w
[1] 1 2 3 4 5
```

外部檔：d:/data2.txt

1
2
3
4
5

□ 其他

read.csv

readLines

read.ssd(foreign) Obtain a Data Frame from a SAS Permanent Dataset, via read.xport

read.xport(foreign) Read a SAS XPORT Format Library

read.xls (gdata): 需安裝perl

ODBC

...

讀寫檔案

■ 內部資料轉至外部資料檔

□ **cat**

```
> cat("==== test ====\n", file="d:/out1.txt")
```

```
> cat(c(1,2,3,4,5), sep="\n", file="d:/out1.txt", append=T)
```

```
d:/out1.txt
==== test ====
1
2
3
4
5
```

□ **write** (向量、matrix、data frame)

```
> x<- matrix(c(1.0, 2.0, 3.0, 4.0, 5.0, 6.0), 2, 3); x
```

```
      [,1] [,2] [,3]
```

```
[1,]  1   3   5
```

```
[2,]  2   4   6
```

```
> write(t(x), ncolumns=3, file="d:/out2.txt")
```

```
> write( x, ncolumns=3, file="d:/out3.txt")
```

```
d:/out2.txt
1 3 5
2 4 6
```

```
d:/out3.txt
1 2 3
4 5 6
```

讀寫檔案

- ***write.table***, ***write.csv*** (matrix 、 data frame)

```
> write.table(x, file="d:/out4.txt", row.names=F, col.names=F, sep=" ")
```

```
> write.table(x, file="d:/out5.txt", row.names=T, col.names=T, sep=" ")
```

```
d:/out5.txt  
"V1" "V2" "V3"  
"1" 1 3 5  
"2" 2 4 6
```

- ***format***

```
> x<- matrix(c(1, 2.2, 3.33, 4.4, 5.55, 6.0), nrow=3, ncol=2)
```

```
> write(format(t(x), nsmall=3), ncolumns=3, file="d:/out6.txt")
```

```
d:/out6.txt  
1.000 4.400 2.200  
5.550 3.330 6.000
```

儲存結果

■ 文字結果的儲存

- copy command上的結果到...

- ***sink***

 - > sink(file="d:/out.txt")

 - > cat(" ===== test =====\n")

 - > y.lm<- lm(Y~X1+X2, data=mydata)

 - > print(summary(y.lm))

 - > cat("dim(mydata)=", paste(dim(mydata), collapse="x",sep=""), "\n")

 - ...

 - > sink() #結束

- ***cat***

- **write.table, write, ...**

儲存結果

■ 圖形結果的儲存

- pdf, png, jpeg, bmp, postscript : 開始

- dev.off : 結束

- 例子

 - > pdf(file="d:/test.pdf") #開啟pdf檔

 - > plot(rnorm(10)) #在dev.off執行前的圖形都會存到test.pdf中

 - ... #圖形不會顯示在螢幕上

 - > dev.off() #關閉pdf檔

自訂函數 (function)

■ 優點

- 避免物件不必要的儲存
 - 儲存物件 <<- 或 assign
 - 名稱 <<- 物件
 - assign(名稱, 物件, pos=1)
 - dput (dget), cat, write, write.table, ...
- 易重複執行
- 遞迴
 - 數值積分

自訂函數 (function)

■ 格式

```
fun_name <- function(...){  
  ...  
}
```

1. 參數(給定default, 則可不輸入)
2. 不給定

□ fun_name: 自行定義

□ 執行: fun_name(...)

■ 提示語">"後編輯與執行

■ 儲存成R code (*.R) -> source (*.R) -> 執行

自訂函數 (function)

■ 例子1

```
> f1<- function(){
  x<- rnorm(100)
  x_sy<- f2(x)
  print(round(x_sy,2))
  assign("x.out", x, pos=1) # x.out<- x
}
> f2<- function(v, na.rm0=F){
  v_mean<- mean(v, na.rm=na.rm0)
  v_qv<- quantile(v, seq(0,1, 0.25), na.rm=na.rm0)
  c(Mean=v_mean, v_qv)
}
> f1()
Mean      0%      25%      50%      75%      100%
-0.12 -2.27 -0.80 -0.12  0.61  2.20
```

自訂函數 (function)

■ 例子2

```
> f3<- function(){  
  x<- matrix(rnorm(100*3), ncol=3)  
  x_sy<- apply(x, 2, FUN=f2, na.rm0=T)  
  print(round(x_sy,3))  
}
```

> f3()

	[,1]	[,2]	[,3]
Mean	0.091	0.025	0.114
0%	-2.377	-2.446	-2.185
25%	-0.552	-0.521	-0.534
50%	0.110	0.118	0.100
75%	0.816	0.784	0.854
100%	2.556	1.926	2.894

自訂函數 (function)

■ 例子3

```
> f2(v=runif(100))
```

Mean	0%	25%	50%	75%	100%
0.4735850	0.0098036	0.2123027	0.4789563	0.6907093	0.9973119

自訂函數 (function)

■ 例子4

將f1, f2存在檔案中

□ File ⇒ Source R code ...

⇒ select d:/test1.R

⇒ > f1()

□ > source("d:/test1.R")

> **f1()#1**

□ File ⇒ Source R code ...

⇒ select d:/test2.R

□ > source("d:/test2.R")

d:/test1.R

```
f1<-function(){  
  ...  
}  
f2<-function(v, na.rm0=F){  
  ...  
}
```

d:/test2.R

```
f1<-function(){  
  ...  
}  
f2<-function(v, na.rm0=F){  
  ...  
}  
f1()
```