

R 基础绘图系统 & 动态报告简介

第五届 R 会议上海会场

邱怡轩 谢益辉 魏太云



2012 年 11 月 4 日

目录

- 1 R 基础图形库
- 2 动画系统
 - animation
 - R2SWF
- 3 knitr 简介
 - 语法概览和全局参数
 - 代码和文本输出
 - 图片和动画
 - 缓存
 - 代码引用和外部代码
 - 编辑器
 - 常见错误
 - 应用案例与学习资源

提纲

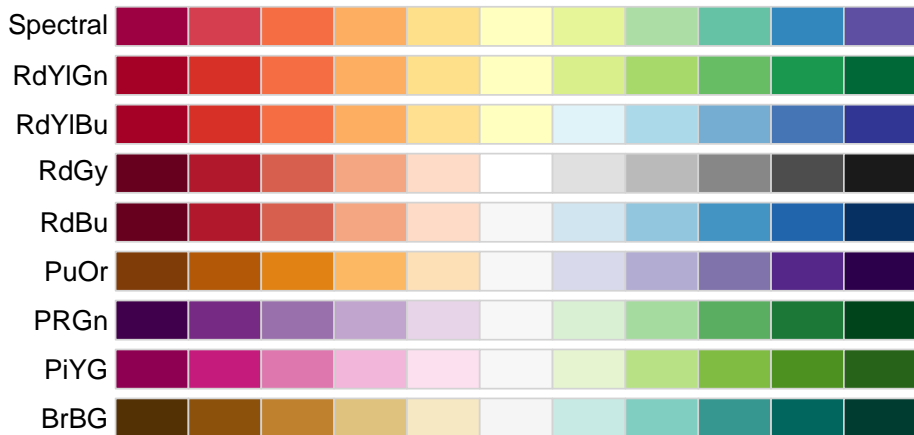


1 R 基础图形库

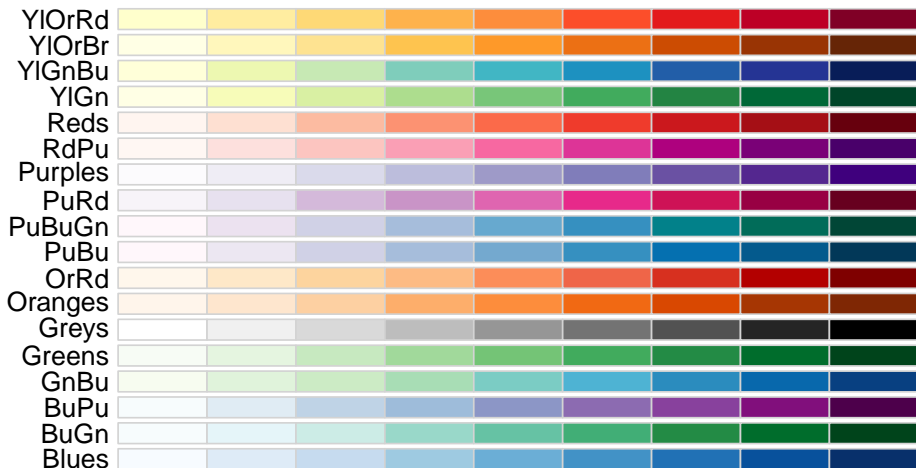
色彩选择: 离散系 (RColorBrewer 包提供)



色彩选择: 双色系 (RColorBrewer 包提供)

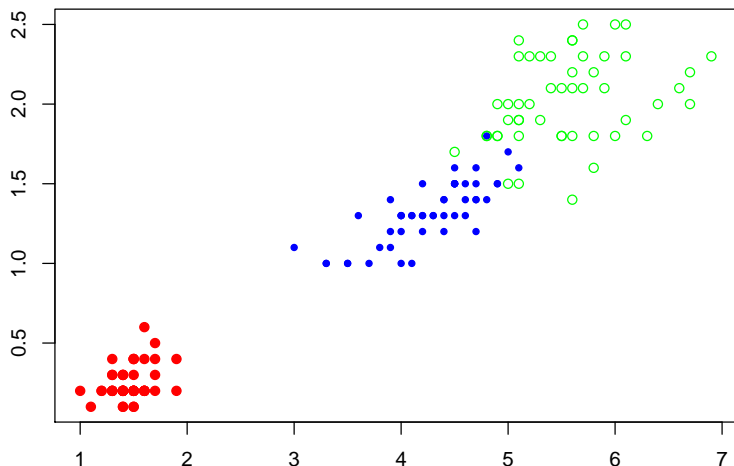


色彩选择: 单色系 (RColorBrewer 包提供)



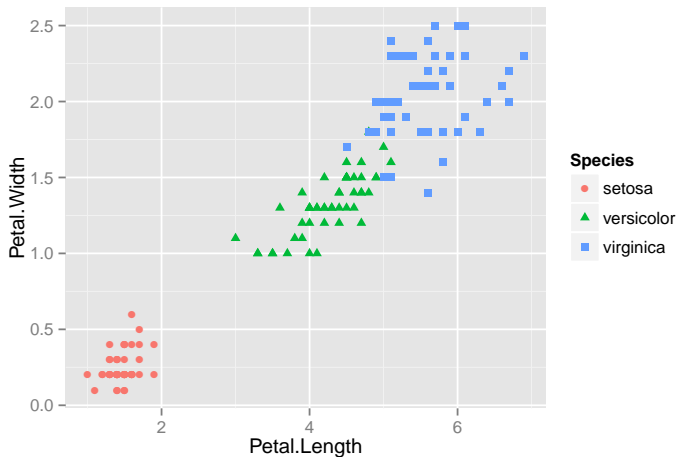
散点图

```
with(iris, plot(Petal.Length, Petal.Width, col = rep(c("red", "blue", "green"),  
  each = 50), pch = rep(19:21, each = 50)))
```

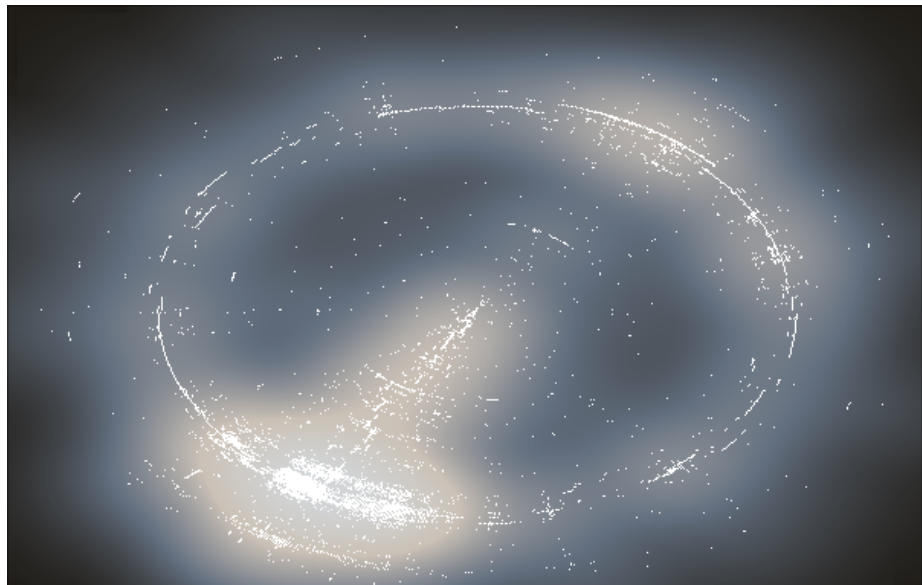


散点图 (ggplot2 版本)

```
qplot(Petal.Length, Petal.Width, color = Species, shape = Species, data = iris)
```

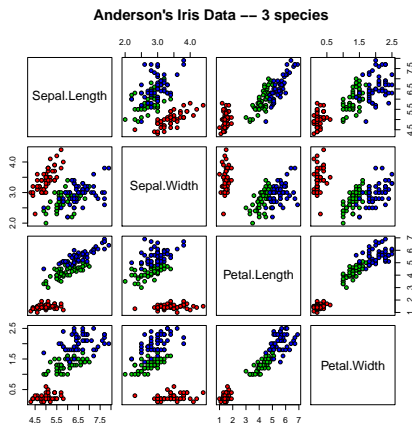


散点图 — 平滑处理：宋词星云



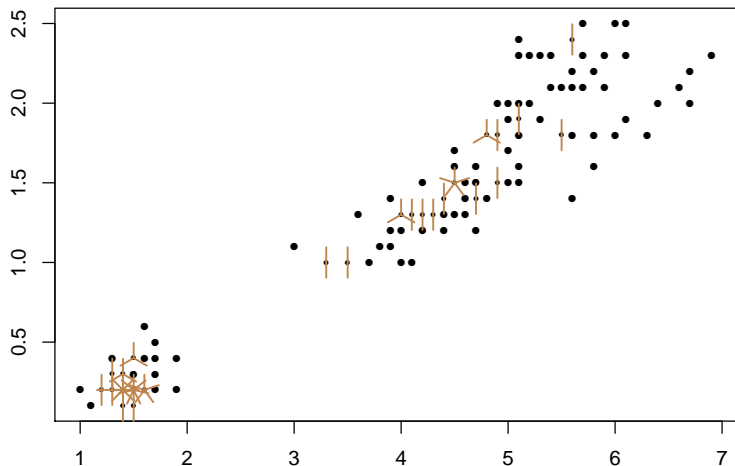
散点图矩阵

```
pairs(iris[1:4], main = "Anderson's Iris Data -- 3 species", pch = 21, bg = c("red",  
  "green3", "blue")[unclass(iris[["Species"]])])
```



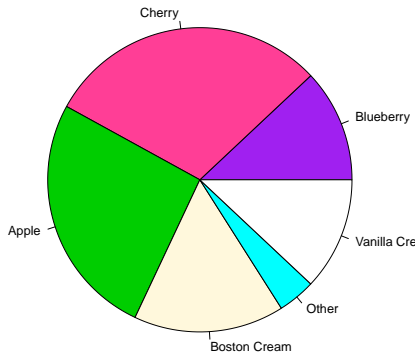
散点图 — 太阳花版本

```
sunflowerplot(iris[, 3:4])
```



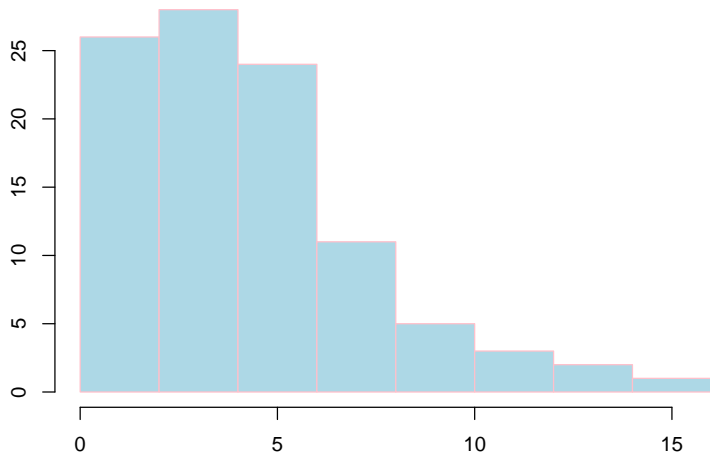
密度图 — 饼图

```
pie.sales <- c(0.12, 0.3, 0.26, 0.16, 0.04, 0.12)
names(pie.sales) <- c("Blueberry", "Cherry", "Apple", "Boston Cream", "Other",
  "Vanilla Cream")
pie(pie.sales, col = c("purple", "violetred1", "green3", "cornsilk", "cyan",
  "white"))
```



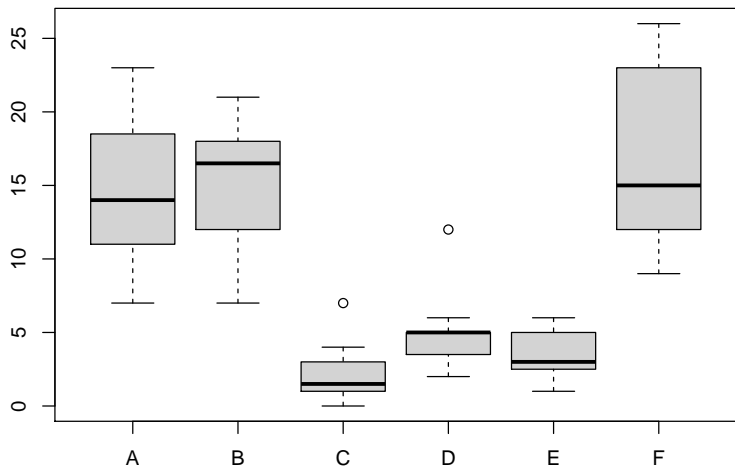
密度图 — 直方图

```
hist(rchisq(100, df = 4), col = "lightblue", border = "pink", main = "")
```



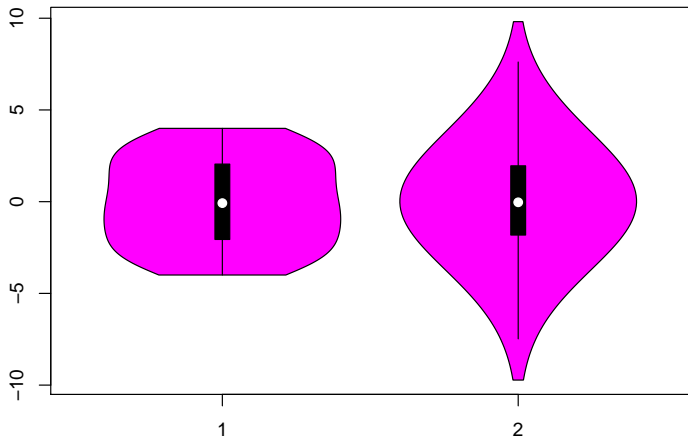
密度图 — 箱线图

```
boxplot(count ~ spray, data = InsectSprays, col = "lightgray")
```



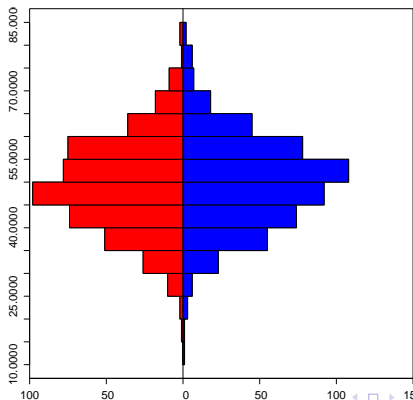
密度图 — 小提琴图 (vioplot 包)

```
library(vioplot)
uniform <- runif(2000, -4, 4)
normal <- rnorm(2000, 0, 3)
vioplot(uniform, normal)
```

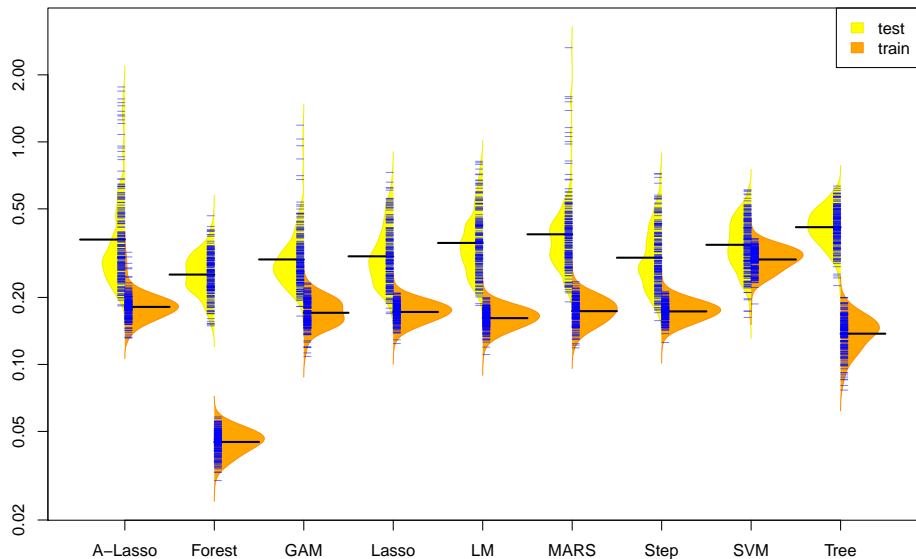


密度图 — 背靠背直方图 (Hmisc 包)

```
library(Hmisc)
age <- rnorm(1000, 50, 10)
sex <- sample(c("female", "male"), 1000, TRUE)
out <- histbackback(split(age, sex))
barplot(-out[["left"]], col = "red", horiz = TRUE, space = 0, add = TRUE, axes = FALSE)
barplot(out[["right"]], col = "blue", horiz = TRUE, space = 0, add = TRUE, axes = FALSE)
```

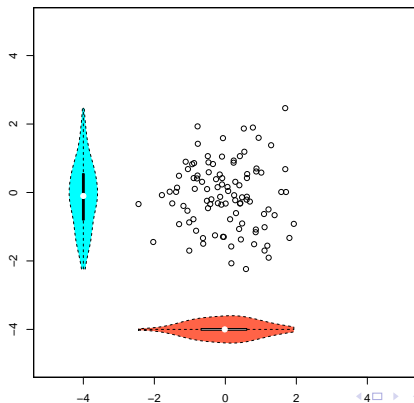


密度图 — beanplot(beanplot 包)



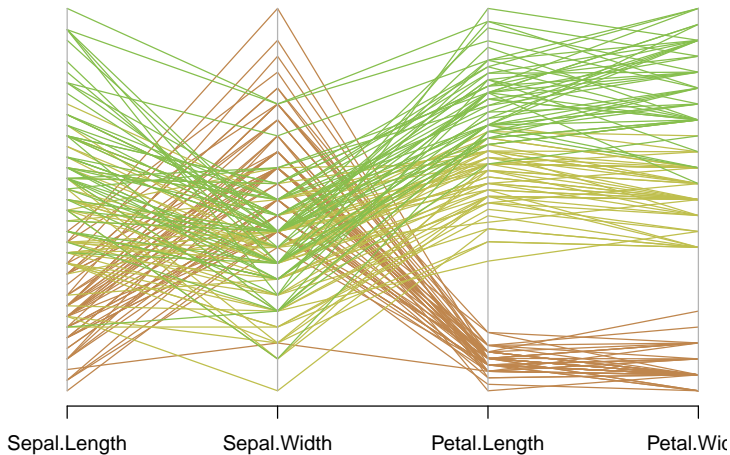
散点图 + 密度图 — 小提琴图

```
x <- rnorm(100)
y <- rnorm(100)
plot(x, y, xlim = c(-5, 5), ylim = c(-5, 5))
vioplot(x, col = "tomato", horizontal = TRUE, at = -4, add = TRUE, lty = 2,
        rectCol = "gray")
vioplot(y, col = "cyan", horizontal = FALSE, at = -4, add = TRUE, lty = 2)
```



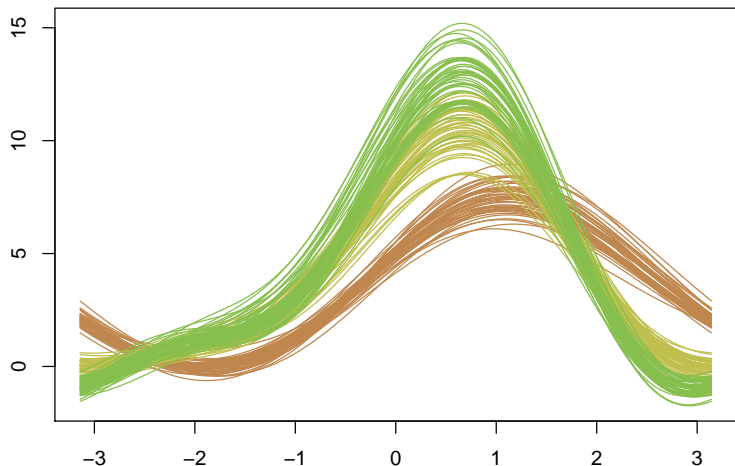
多元图 — 平行坐标图 (MASS 包)

```
library(MASS)
parcoord(iris[, 1:4], col = rep(2:4, each = 50))
```



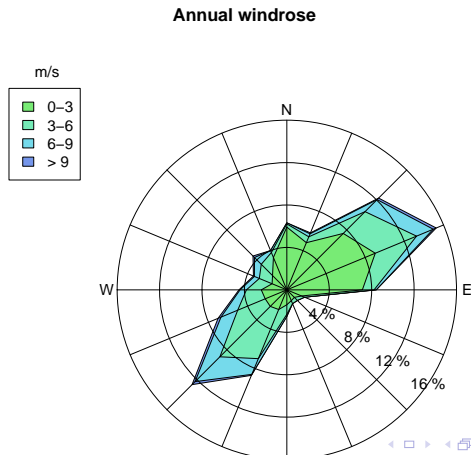
多元图 — 调和曲线图 (MSG 包)

```
library(MSG)  
andrews_curve(iris[, 1:4], col = rep(2:4, each = 50))
```



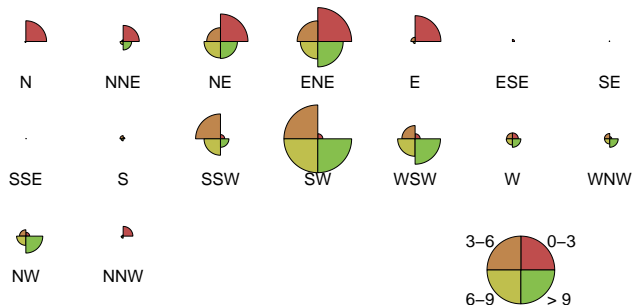
多元图 — 蛛网图 (climatol 包)

```
library(climatol)
data("windfr")
rosavent(windfr, 4, 4, ang = -3 * pi/16, main = "Annual windrose")
```



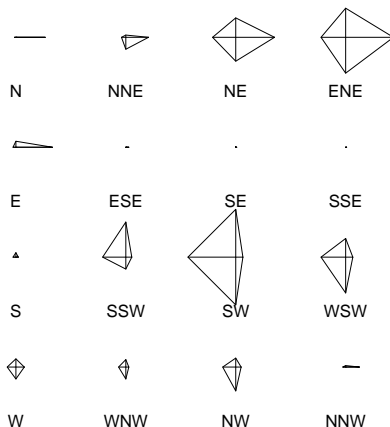
多元图 — 雷达图

```
palette(rainbow(12, s = 0.6, v = 0.75))
stars(t(windfr), len = 0.8, key.loc = c(14, 1.5), ncol = 7, main = "", draw.segments = TR
```



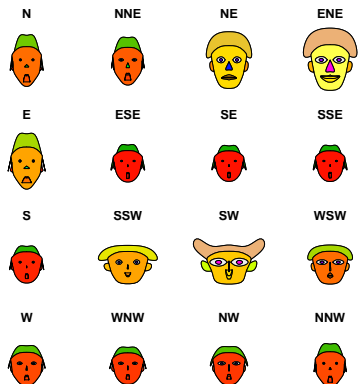
多元图 — 星状图

```
stars(t(windfr))
```



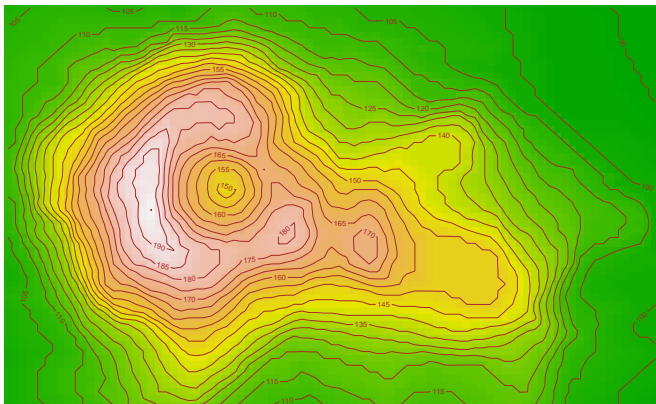
多元图 — 脸图 (aplpack)

```
aplpack::faces(t(windfr))
```



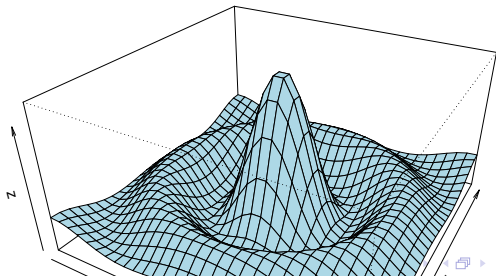
地理图 — 等高线图

```
x <- 10 * (1:nrow(volcano))
x.at <- seq(100, 800, by = 100)
y <- 10 * (1:ncol(volcano))
y.at <- seq(100, 600, by = 100)
image(x, y, volcano, col = terrain.colors(100), axes = FALSE)
contour(x, y, volcano, levels = seq(90, 200, by = 5), add = TRUE, col = "brown")
```



三维表面图

```
x <- seq(-10, 10, length = 30)
y <- x
f <- function(x, y) {
  r <- sqrt(x^2 + y^2)
  10 * sin(r)/r
}
z <- outer(x, y, f)
z[is.na(z)] <- 1
persp(x, y, z, theta = 30, phi = 30, expand = 0.5, col = "lightblue")
```



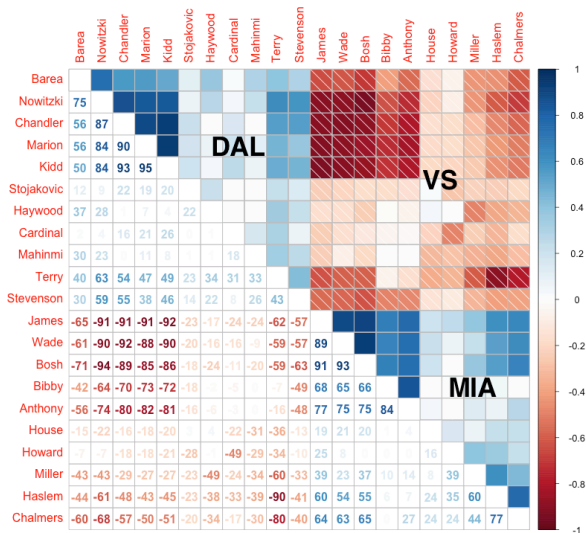
三维图 (演示 rgl 包)

```
library(rgl)
x = y = seq(1, 5, 0.1)
m = outer(x, y, function(a, b) beta(a, b))
persp3d(x, y, m, col = "green3", zlab = "Beta(x, y)")
```

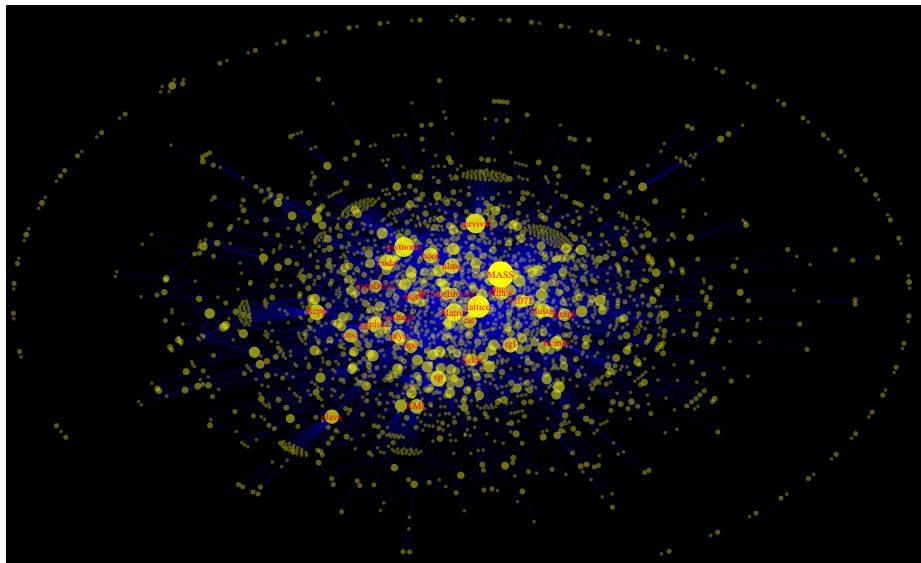
标签云图 (wordcloud 包)

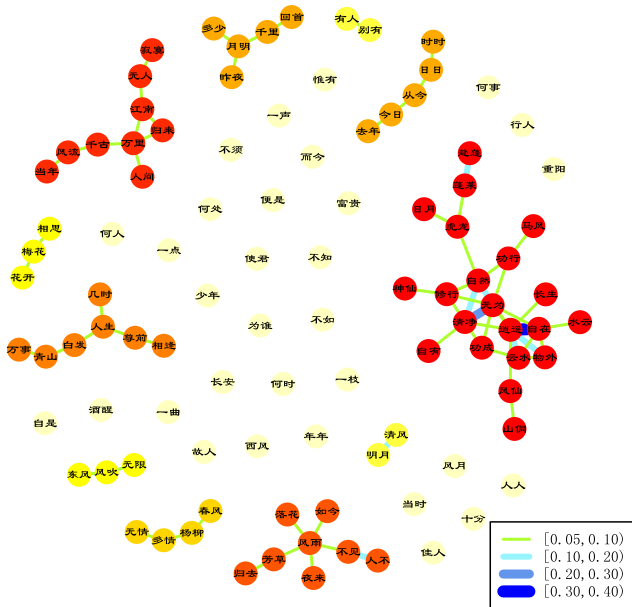


关系矩阵图 (corrplot 包): 来自 <http://www.d3coder.com>



网络图 (igraph 包)





提纲



- ## 2 动画系统
- animation
 - R2SWF

animation 包

- ▶ 内置了许多统计模型的动画展示

```
knn.ani() kmeans.ani() buffon.needle() ...
```

- ▶ 提供了生成动画的统一框架

```
ani.options() saveSWF() saveMovie() saveHTML() ...
```

生成动画文件

▶ 将 R 代码生成的图片保存为 SWF 文件

- ▶ 需要下载 SWFTools (<http://www.swftools.org/>) 到 swftools 参数指定的路径中

```
saveSWF({ par(mar = c(3, 3, 1, 1.5), mgp = c(1.5, 0.5, 0)) knn.ani(test =
matrix(rnorm(16), ncol = 2), cl.pch = c(16, 2)) }, swf.name = "kNN.swf",
interval = 1.5, nmax = 40, swftools = "d:/tmp")
```

▶ 保存为 GIF 动画

- ▶ 需要下载 ImageMagick (<http://www.imagemagick.org/>)

```
ani.options(convert = "d:/tmp/convert.exe") saveGIF({ for (i in 1:10)
plot(runif(10), ylim = 0:1) })
```

R2SWF 包

- ▶ 将图片转换成 SWF 格式，无需安装或下载其它的软件
- ▶ `image2swf()` 用于转换 PNG 或 JPG 图片
- ▶ `svg2swf()` 用于转换 SVG 图片（从而生成的 SWF 动画是矢量的）
- ▶ `dev2swf()` 自动转换 R 代码生成的图片
- ▶ `swf2html()` 用于将 SWF 动画在网页中呈现

例子

► 生成一段布朗运动的动画

```
if (capabilities("cairo")) {  
  olddir = setwd(tempdir())  
  svg("Rplot%03d.svg", onefile = FALSE)  
  set.seed(123)  
  x = rnorm(5)  
  y = rnorm(5)  
  for (i in 1:100) {  
    plot(x <- x + 0.1 * rnorm(5), y <- y + 0.1 * rnorm(5), xlim = c(-3,  
      3), ylim = c(-3, 3), col = "steelblue", pch = 16, cex = 2, xlab = "x",  
      ylab = "y")  
  }  
  dev.off()  
  output = svg2swf(sprintf("Rplot%03d.svg", 1:100), interval = 0.1)  
  swf2html(output)  
  setwd(olddir)  
}
```

例子

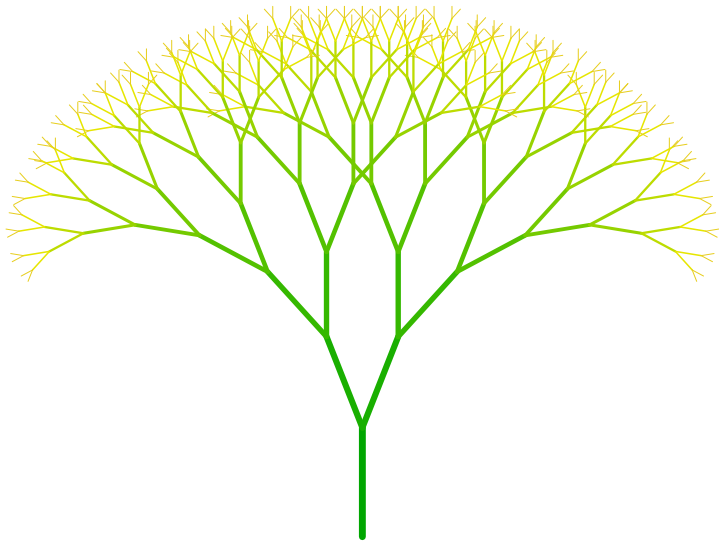
► 演示 Kmeans 聚类

```
library(animation)
output = dev2swf({
  par(mar = c(3, 3, 1, 1.5), mgp = c(1.5, 0.5, 0))
  kmeans.ani()
}, output = "test.swf")
swf2html(output)
```

网页中的动画

- ▶ flash 形式
- ▶ gif 形式
- ▶ js 形式: 观看例子<http://taiyun.github.com/>

PDF 中的动画



分形树代码

```
drawTree <- function(x1, y1, angle, n, lwd, len, col) {  
  if (n >= 1) {  
    x2 <- x1 + cos(angle) * len[n] * 10  
    y2 <- y1 + sin(angle) * len[n] * 10  
    lines(c(x1, x2), c(y1, y2), lwd = lwd[n], col = col[n])  
    drawTree(x2, y2, angle - pi/9, n - 1, lwd, len, col)  
    drawTree(x2, y2, angle + pi/9, n - 1, lwd, len, col)  
  }  
}  
  
COL <- rev(terrain.colors(16)[1:9])  
windows(width = 6, height = 4)  
par(mar = c(0, 0, 0, 0), ask = TRUE)  
for (i in 1:9) {  
  plot(0, xlim = c(-260, 260), ylim = c(0, 450), type = "n", axes = F)  
  drawTree(0, 0, angle = pi/2, n = i, lwd = (10 - i):9, len = (10 - i):9,  
    col = COL[(10 - i):9])  
}
```

提纲

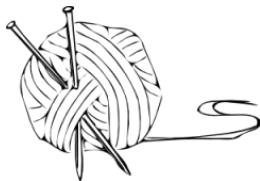
3 knitr 简介

- 语法概览和全局参数
- 代码和文本输出
- 图片和动画
- 缓存
- 代码引用和外部代码
- 编辑器
- 常见错误
- 应用案例与学习资源



名称

- ▶ $\text{knitr} = \text{knit} + \text{R}$ (类似 $\text{S} + \text{weave}$);
- ▶ 发音: neater?
- ▶ R 小写
- ▶ Logo:



语法概览

► Rnw 文件:

段落: `<<name, echo=TRUE, tidy=TRUE>>=`
这里是段落代码
`@`

行内: `\Sexpr{这里是行内代码}`

► Rmd 文件:

段落: ```` {r test-r, engine='R', comment=NA, dev='CairoPNG'}`
这里是段落代码
`````

行内: ``r 这里是行内代码``

## ► Rhtml 文件:

段落: `<!--begin.rcode my-label, fig.width=5, dev=png`  
这里是段落代码  
`end.rcode-->`

行内: `<!--rinline 这里是行内代码 -->`

# 全局参数

## ► Rnw 文件:

```
\SweaveOpts{fig.align='center', cache=TRUE, tidy=FALSE, comment=NA}
```

或:

```
<<setup, include=FALSE, cache=FALSE>>=
opts_chunk$set(fig.align='center', cache=TRUE, tidy=FALSE, comment=NA)
@
```

## ► Rmd 文件:

```
`r opts_chunk$set(fig.width=6, fig.height=6, fig.path='')`
```

## ► Rhtml 文件:

```
<!--begin.rcode setup,echo=FALSE,results=hide,message=FALSE,cache=FALSE
 opts_chunk$set(fig.width=5, fig.height=5, ,fig.align='center')
end.rcode-->
```

# 段落代码和文本输出参数

- ▶ **eval**: (**TRUE**; 逻辑): 是否执行代码
- ▶ **tidy**: (**TRUE**; 逻辑): 是否整理代码
- ▶ **prompt**: (**FALSE**; 逻辑): 是否添加引导符 '>'
- ▶ **highlight**: (**TRUE**; 逻辑): 是否高亮代码
- ▶ **size**: ('normalsize'; 字符): 字体大小
- ▶ **background**: ('#F7F7F7'; 字符或数值): 背景颜色
- ▶ **comment**: ('##'; 字符): 结果输出前缀符号
- ▶ **echo**: (**TRUE**; 逻辑或数值): 是否输出代码或输出哪些行
- ▶ **results**: ('markup'; 字符): 装裱输出 ('markup')、原样输出 ('asis')、隐藏 ('hide')
- ▶ **warning, error, message**: (**TRUE**; 逻辑): 是否显示相应信息
- ▶ **split**: (**FALSE**; 逻辑): 是否剥离代码和文本到外部文件
- ▶ **include**: (**TRUE**; 逻辑): 是否保留代码或结果到最终文档

# 代码

```
<<example-sexpr, prompt=TRUE>>=
```

```
x = date()
```

```
y = 1:10
```

```
matrix(1:6, nrow=2)
```

```
@
```

本幻灯片最后修改时间是\Sexpr{x}, y值是\Sexpr{y}。

```
> x = date()
```

```
> y = 1:10
```

```
> matrix(1:6, nrow = 2)
```

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

本幻灯片最后修改时间是 Sun Nov 04 14:12:47 2012, y 值是 1, 2, 3, 4, 5, 6, 7, 8, 9, 10。

# 主要参数

- ▶ `fig.path`: ('figure/'; 字符): 图片路径, 支持前缀模式 ('figure/prefix-')
- ▶ `fig.keep`: ('high'; 字符): 保存图形类型, 高级图形 ('high')、不保存 ('none')、所有图形 ('all')、第一张 ('first')、最后一张 ('last')
- ▶ `fig.show`: ('asis'; 字符): 展示方式, 紧随代码输出 ('asis')、最后统一输出 ('hold')、动画输出 ('animate')
- ▶ `dev`: (LaTeX 为 'pdf', HTML/markdown 为 'png'; 字符): 输出设备, knitr 支持很多种设备
- ▶ `fig.width`, `fig.height`: (7; 数值): 图片文件的宽、高 (英寸 2.54cm 为单位)
- ▶ `out.width`, `out.height`: (NULL; 字符): 图片在输出文档中的宽、高
- ▶ `fig.align`: ('default'; 字符): 对齐方式, 不做调节 ('default')、左 ('left')、右 ('right')、居中 ('center')
- ▶ `interval`: (1; 数值): 动画参数, 切换画面时间, 单位为秒



# 图片和动画

## ▶ 图片

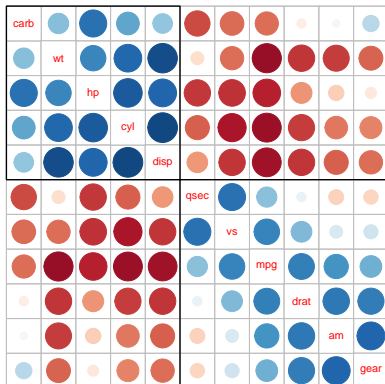
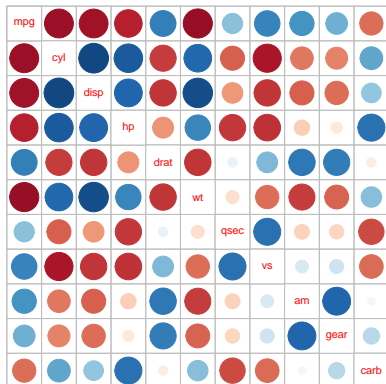
```
<<fig.width=6.2, fig.height=6.2, out.width='0.46\\linewidth'>>=
require(corrplot)
corrplot(cor(mtcars), tl.pos="d", cl.pos="no")
corrplot(cor(mtcars), order="h", tl.pos="d", cl.pos="no", addrect=2)
@
```

## ▶ 动画

```
<<fig.width=3, fig.height=3, out.width='.35\\linewidth', fig.show='animate'>>=
par(mar = rep(2.3, 4))
for (i in seq(pi/2, -4/3 * pi, length = 12)) {
 plot(0, 0, pch = 20, ann = FALSE, axes = FALSE)
 arrows(0, 0, cos(i), sin(i))
 axis(1, 0, "6"); axis(2, 0, "9")
 axis(3, 0, "12"); axis(4, 0, "3"); box()
}
@
```

# 图片输出

```
require(corrplot)
corrplot(cor(mtcars), tl.pos = "d", cl.pos = "n")
corrplot(cor(mtcars), order = "h", tl.pos = "d", cl.pos = "n", addrect = 2)
```



# 动画输出

```
par(mar = rep(2.3, 4))
for (i in seq(pi/2, -4/3 * pi, length = 12)) {
 plot(0, 0, pch = 20, ann = FALSE, axes = FALSE)
 arrows(0, 0, cos(i), sin(i))
 axis(1, 0, "6"); axis(2, 0, "9")
 axis(3, 0, "12"); axis(4, 0, "3"); box()
}
```

# 缓存

- ▶ 为什么需要缓存?
- ▶ 主要参数:
  - ▶ `cache`: (`FALSE`; 逻辑): 是否开启缓存
  - ▶ `cache.path`: (`'cache/'`; 字符): 缓存路径

# 代码引用 (对 chunk1)

## ① <<>> 格式 (可以多级别引用):

```
<<chunk2>>=
<<chunk1>>
@
```

## ② 相同标签模式 (后一个代码必须置空):

```
<<chunk1, echo=FALSE, results='markup'>>=
@
```

## ③ 使用ref.label 参数 (这样可以分离代码和结果):

```
<<chunk2, ref.label='chunk1', echo=FALSE, results='markup'>>=
@
```

## ④ 使用run\_chunk() 函数 (支持嵌套结构)

```
<<a>>=
x = 1; x
run_chunk('b'); x
@
<>=
x = 2
@
<<c>>=
run_chunk('a'); x
@
```

# 外部代码

## ▶ 外部代码out.R:

```
@knitr Q1
rnorm(10)
@knitr Q2
sample(10)
```

## ▶ 引用外部代码

```
<<set-options, echo=FALSE, cache=FALSE>>=
read_chunk('out.R')
@
<<Q1>>=
@
<<Q2>>=
@
```

## ▶ 结果

```
> rnorm(10)

[1] -0.4126 1.1318 -0.3002 1.5228 -0.7210 -0.3329 -1.0284 0.2687
[9] -0.4417 0.5721
```

```
> sample(10)

[1] 2 4 8 9 5 6 7 3 10 1
```

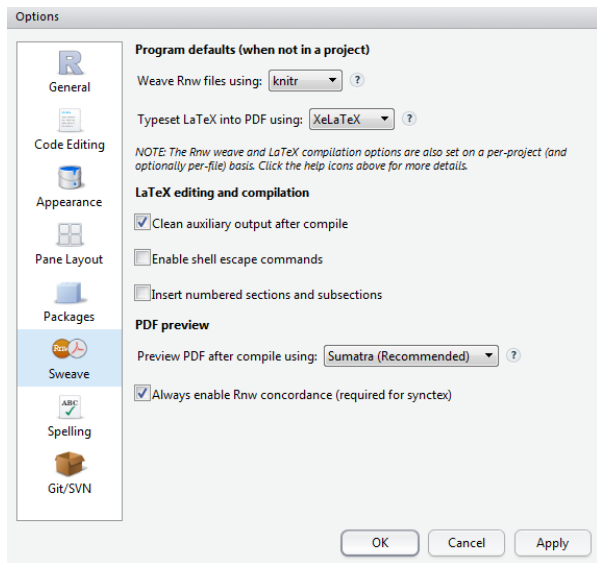
# 编辑器的配置

- ▶ **RStudio**
- ▶ **LyX**
- ▶ Emacs Org-mode
- ▶ Eclipse
- ▶ TeXStudio
- ▶ WinEdt

参见：

- ▶ <http://yihui.name/knitr/demo/rstudio/>
- ▶ <http://yihui.name/knitr/demo/lyx/>
- ▶ <http://yihui.name/knitr/demo/editors/>

# RStudio(10 月 31 日发布最新版 0.97.168)





# 常见错误

## ▶ 参数问题

label参数重复  
 echo参数越界  
 图片大小设置不当，比如混淆fig.width和out.width

## ▶ 中文问题：UTF-8 编码，代码不能有中文变量；最好 X<sub>Y</sub>TEX 编译

```
Rprofile.site 文件配置
options(stringsAsFactors = FALSE, help_type = "html", encoding = "UTF-8")
Sys.setlocale("LC_CTYPE", "chs")
```

## ▶ Beamer 中的问题

[fragile] 参数问题：  
`\begin{frame}[fragile]{代码例子}` 含代码的片子中勿忘fragile参数  
`\AtBeginSection[]` 等配置环境中，{frame}后不能加fragile参数

# 应用案例

- ▶ ggplot2 0.9.0 guide 文档
- ▶ ggbio 文档: <http://tengfei.github.com/ggbio/>
- ▶ animation: <http://taiyun.cos.name>
- ▶ vistat: <http://vis.supstat.com/>
- ▶ 更多 (包括书籍、网站等):  
<http://yihui.name/knitr/demo/showcase/>

# 在线报告

- ▶ <http://public.opencpu.org/apps/knitr/>
- ▶ <http://r.psyapp.com/apps/knitr/>
- ▶ Rpubs: <http://www.rpubs.com/>

# 学习资源

- ▶ 主页: <http://yihui.name/knitr/>
- ▶ 忍者秘笈: <http://t.cn/z09tnqK>
- ▶ 演示文件: <http://yihui.name/knitr/demo/>
- ▶ 开发页面: <https://github.com/yihui/knitr/>
- ▶ 邮件列表: <http://groups.google.com/group/knitr>
- ▶ 文档下载: <https://github.com/yihui/knitr/downloads>

# 谢谢大家！



邱怡轩: `yixuan.qiu@cos.name`

谢益辉: `yihui.xie@cos.name`

魏太云: `taiyun.wei@cos.name`

## Q/A?