R 语言学习笔记

作者：Jack Zhou

QQ:10853**172

日期：2019-08-10

目录

一.R 语言学习网站

1. https://www.r-project.org/

   R-CRAN Comprehensive R Archive Network（R 综合典藏网）

2. https://github.com/search?q=R

   托管在 github 上的 R 语言包

3. https://r-forge.r-project.org/

   一部分 R 包的托管

4. https://www.bioconductor.org/

   生物信息 R 包的托管

5. https://www.rmetrics.org/

   金融领域 R 包的扩展

二.几个综合性网站

6. https://www.rdocumentation.org/

   install.packages("Rdocumentation")可以扩充 R 语言基础帮助系统

7. https://rdrr.io/（可以云上 run jypter）

   这个网站包含的 R packages 介绍很完善

三.如何加快 R 语言运行速度

Microsoft R Open is the enhanced distribution of R from Microsoft Corporation. It is a complete open source platform for statistical analysis and data science.

The current version, Microsoft R Open 3.5.3, is based on (and 100% compatible with) R-3.5.3, the most widely used statistics software in the world, and is therefore fully compatibility with all packages, scripts and applications that work with that version of R. It includes additional capabilities for improved performance, reproducibility, as well as support for Windows and Linux-based platforms.

Like R, Microsoft R Open is open source and free to download, use, and share.

网站：https://mran.microsoft.com/open

下面用一个例子对比下运行效率

测试代码

```
list1<-list()
time1<-Sys.time()
for(i in 1:10){
    x<-matrix(rnorm(100000),1000,1000)
    list1[[i]]<-x%*%x%*%x%*%x%*%x
}
time2<-Sys.time()
cost<-(time2-time1)/100
cost
```

1. 环境　CRAN 版本

```
R version 3.6.0 (2019-04-26) -- "Planting of a Tree"
Copyright (C) 2019 The R Foundation for Statistical Comput
ing
Platform: x86_64-w64-mingw32/x64 (64-bit)
```

```
> list1<-list()
> time1<-Sys.time()
> for(i in 1:10){
+   x<-matrix(rnorm(100000),1000,1000)
+   list1[[i]]<-x%*%x%*%x%*%x%*%x
+ }
> time2<-Sys.time()
> cost<-(time2-time1)/100
> cost
Time difference of 0.0125295 mins
```

运行时间：**0.0125295** 分钟

2. 环境 Microsoft R Open 3.5.3

```
Microsoft R Open 3.5.3
The enhanced R distribution from Microsoft
Microsoft packages Copyright (C) 2019 Microsoft Corporation
> list1<-list()
> time1<-Sys.time()
> for(i in 1:10){
+   x<-matrix(rnorm(100000),1000,1000)
+   list1[[i]]<-x%*%x%*%x%*%x%*%x
+ }
> time2<-Sys.time()
> cost<-(time2-time1)/100
> cost
Time difference of 0.05329 secs
```

运行时间：**0.05329** 秒

效率比值

```
0.0125295*60/0.05329
>[1] 14.10715
```

通过一个简单的例子可以发现保守估计有 **1** 个数量级的差距。

微软官网介绍 R 多线程的优点

The Benefits of Multithreaded Performance with Microsoft R Open

连接地址：**https://mran.microsoft.com/documents/rro/multithread#mt-mac**

四.两个关于 R 语言的期刊

1. The R Journal

https://journal.r-project.org

2. Journal of Statistical Software

https://www.jstatsoft.org