R与MALTAB及VBA混合编程

中国人民大学　李青龙

Why

R是一款功能强大的能够进行统计研究、数据分析的得力工具。不仅免费、开源，而且算法更新速度快。

Excel是最为广泛使用的交互式表格，在商务办公领域有着无可比拟的市场占有率。

VBA是Excel高阶使用者的得力助手，能够批量化的操作提供工作效率。

MATLAB是由MathWorks公司推出的科学计算语言，其向量化的思想与R相同。

Why R

Why Excel

Why VBA

Why MATLAB

混合编程

每种工具各有所长通过混合编程进行交互显得尤为重要。

搭配种类繁多，而内容又不复杂，几乎找不到制定的书籍。

通过搜索等方式学习混合编程需要很大的学习成本。

在讲座中很容易当场消化理解。

接口、Server、插件

创建对象

建立连接

数据及变量传递

代码传递运行

断开连接

背景介绍

网址： http://rcom.univie.ac.at/

Statconn Project

他们所做的项目不仅仅局限于
RExcel，而是提供各种工具使得R
可以与其他工具混合使用。Rexcel
只是这些工具之一。

statconnDCOM server

statconnWS SOAP/http web service

rcom R package

RExcel

SWord

Use R!

Richard M. Heiberger
Erich Neuwirth

R Through Excel

A Spreadsheet Interface for Statistics,
Data Analysis, and Graphics
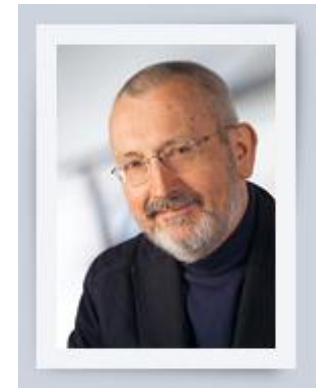
Springer
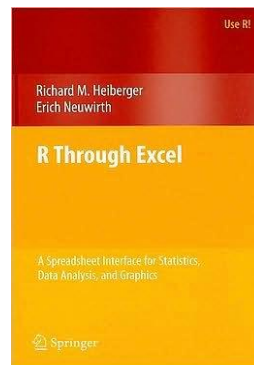
Thomas Baier

Erich Neuwirth

IStatConnector

何为COM？

——The Component Object Model
组件对象模型

statconnDCOM Server 相当于一个软件服务端，可以使R与其他软件融合。

rcom是一个R包，可以使R与其他应用程序同时、交互运行

RExcel 是一个利用用前两者使Excel能够调用R 的加载项

建立COM对象

初始化R

SetSymbol

Evaluate,EvaluateNo Return

GetSymbol

关闭R

释放COM对象

R与MATLAB

MATLAB DEMO

```
R_Obj=actxserver('StatConnectorSrv.StatConnector');

R_Obj.Init('R');

R_Obj.SetSymbol('e',rand([1,100]));
R_Obj.SetSymbol('x',1:100);

R_Obj.EvaluateNoReturn('y <- 3*x+e')
R_Obj.EvaluateNoReturn('lm.sol <-lm(y~x)')
R_Obj.EvaluateNoReturn('z<-lm.sol$coef')

MATLABcoef=R_Obj.GetSymbol('z')

R_Obj.Evaluate('as.vector(lm.sol$coef)')

R_Obj.Evaluate('f<-is.vector(x)')    %FALSE
R_Obj.Evaluate('f<-is.array(x)')     %TRUE
```

- ```
  R Obj=actxserver('StatConnectorSrv.StatConnector');
  ```

- ```
  R_Obj.Init('R');
  ```

- ```
  R_Obj.SetSymbol('e',rand([1,10]));
  ```
- ```
  R_Obj.SetSymbol('x',1:10);
  ```

- *R_Obj.EvaluateNoReturn('x<-as.vector(x)')*
- *R_Obj.EvaluateNoReturn('e<-as.vector(e)')*

- ```
  R_Obj.EvaluateNoReturn('y <- 3*x+e')
  ```
- ```
  R_Obj.EvaluateNoReturn('lm.sol <-lm(y~x)')
  ```
- ```
  R_Obj.EvaluateNoReturn('z<-lm.sol$coef')
  ```

- ```
  MATLABcoef=R_Obj.GetSymbol('z')
  ```
- ```
  MATLAB_Coef=R_Obj.Evaluate('lm.sol$coef')
  ```

- ```
  R_Obj.Close
  ```

MATLAB的R-link

MathWorks公司在2004年的时候就推出了R-link这一工具，提供m文件的免费下载。R-link包括以下函数：

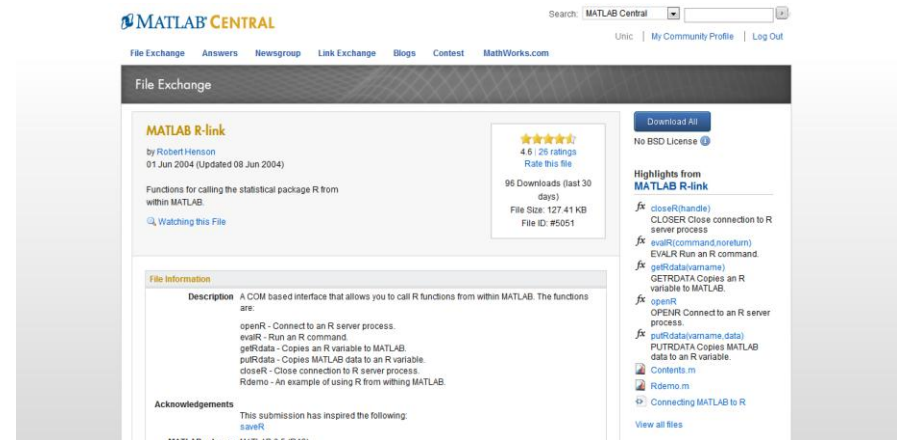**openR** - Connect to an R server process.
evalR - Run an R command.
getRdata - Copies an R variable to MATLAB.
putRdata - Copies MATLAB data to an R variable.
closeR - Close connection to R server process.

个人觉得除了可以使MATLAB编程中代码短一点，并没有必要使用这个工具，它只是把对COM对象的操作写到了函数里而已。

- 网址：
- http://www.mathworks.com/matlabcentral/fileexchange/5051-matlab-r-link
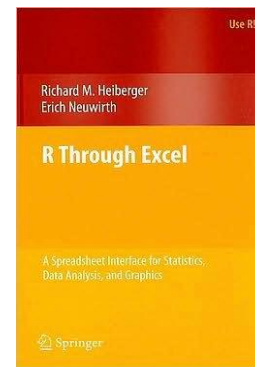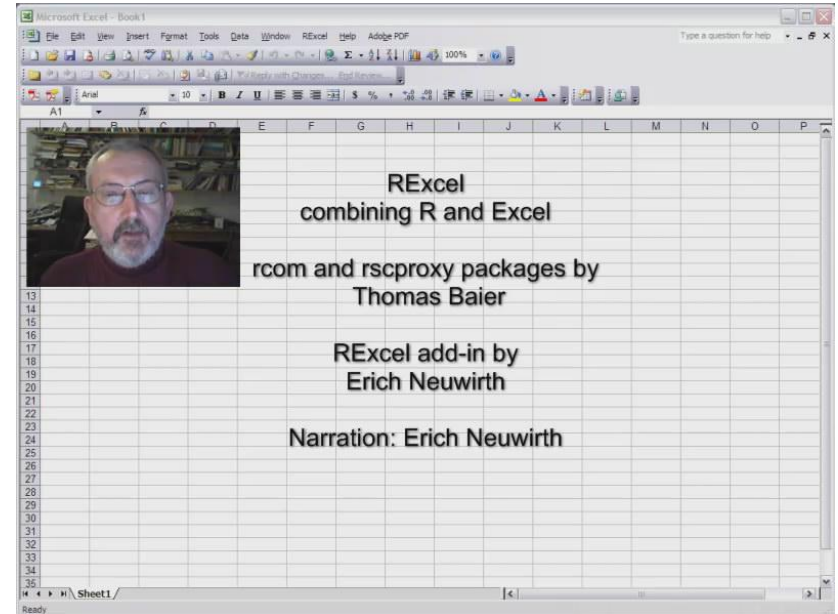
R 与 EXCEL VBA

R与Excel

R与Excel的交互可以说是statconn最为重视的部分，因为是最容易被广泛使用的工具。官网上提供了一个Demo的教学视频，可以保存成MP4的文件自己慢慢学习。

R through Excel是一本专门讲解RExcel的书，不过没有VBA部分的讲解。

在useR的会议slides里面也可以下载到一些有用的资料。

由于两位作者经常从事Rexcel及statconn的收费培训工作，所以很多资料无法得到。

准备工作

为了使用Rexcel，需要在加载项中价值Rexcel，会看到Rexcel的插件菜单。可以实现Spreadsheet与R的交互使用。

如果使用VBA，在VBE编辑器里直接拖动RExcelVBALib即可实现对RExcel2007.xlam的引用。
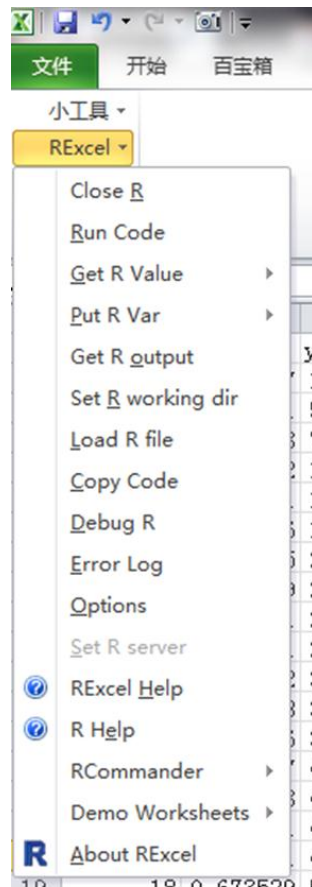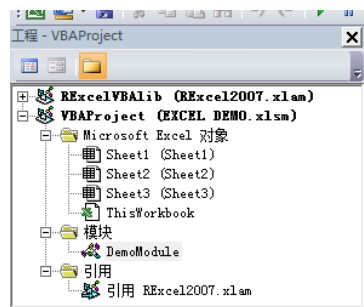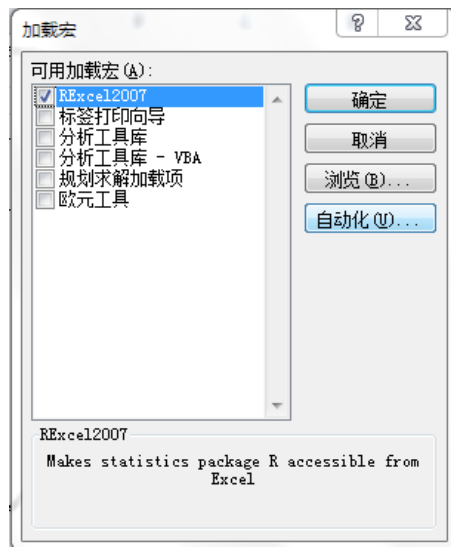
R与VBA

```vba
Public Sub StartRServer(Optional WarningMsgs As Boolean = False)


Public Sub StopRServer(Optional suppressWarnings As Boolean = False)


Public Sub RRun(pString As String, Optional RcmdrExcelMenuUpdate As Boolean = False)


Public Sub PutArray(ArrName As String, ByVal SheetRange As Range, Optional WithRowNames As Boolean = False, Optional WithColNames As Boolean = False)


Public Sub GetArray(RExpression As String, SheetRange As Range, Optional CheckForEmpty As Boolean = False, Optional SaveUndo As Boolean = False, Optional SizeWarning As Boolean = False, Optional WithRowNames As Boolean = False, Optional WithColNames As Boolean = False)


Public Sub RunRFile(RFileName As String)


Sub ExcelGetFocus()
```

```vba
Public Sub PutDataframe(DFName As String, SheetRangeOrName As Variant, Optional WithRowNames As Boolean = False Optional RespectHidden As Boolean = True, Optional CheckDataframeName As Boolean = True, Optional UseRawDataTransfer As Boolean = False)


Public Sub GetDataframe(DFName As String, SheetRange As Range, Optional WithRowNames As Boolean = False, Optional CheckForEmpty As Boolean = False, Optional SaveUndo As Boolean = False, Optional SizeWarning As Boolean = False)


Sub RunRCodeFromRange(commandRange As Range)


Sub InsertCurrentRPlot(Location As Range, Optional picName As String, Optional widthRescale As Double = -1, Optional heightRescale As Double = -1, Optional closeRGraph As Boolean = False)


Sub GetRApply(RFunction As String, trgtRange As Range, ParamArray args() As Variant)


Sub RunRCall(RProc As String, ParamArray args() As Variant)
```

VBA DEMO

```vba
Option Explicit

Sub Demo()

    Worksheets(1).Activate

    Rinterface.StartRServer

    Rinterface.RRun "e<-rnorm(100)"
    Rinterface.RRun "x<-c(1:100)"
    Rinterface.RRun "y = 3*x+e"

    Rinterface.GetArray "x", Range("A2")
    Rinterface.GetArray "e", Range("B2")
    Rinterface.GetArray "y", Range("C2")

    Rinterface.GetRApply "function()coef(lm(y~x))", Range("F2")

    Rinterface.RunRCall "function(x,y)plot(x,y,type=""o"")", Range("A2:A101"), Range("C2:C101")
    Rinterface.InsertCurrentRPlot Range("E5"), widthrescale:=0.5, heightrescale:=0.5, closergraph:=True

    'Rinterface.StopRServer

End Sub
```

谢谢