

GoMRcpp

Descrição:

GoMRcpp.R é um arquivo R-Script desenvolvido por Julimar Pinto e André Caetano no Programa de Pós-Graduação em Ciências Sociais da Pontifícia Universidade Católica de Minas Gerais no ano de 2013. O objetivo de GoMRcpp.R é a modelagem matemática de Grade of Membership (GoM) para dados discretos.

Considera-se que GoMRcpp.R é uma versão atualizada de GOM3.EXE (versão 3.4). GOM3.EXE foi desenvolvido por Peter Charpentier no Departamento de Epidemiology and Public Health da Yale School of Medicine no ano de 1992. O arquivo executável GOM3.EXE, que opera em ambiente MS-DOS, encontra-se disponível para download em: <<http://lib.stat.cmu.edu/DOS/general/.index.html>>. Assim como na versão de Peter Charpentier, as funções matemáticas utilizadas em GoMRcpp.R foram originalmente descritas por Woodbury and Clive (1974).

GoMRcpp.R foi escrito nas linguagens R e C/C++. Para que GoMRcpp.R seja utilizado no sistema operacional Windows da Microsoft é necessário que seja executado em ambiente R com a devida instalação do Rtools e, pelo menos, com a instalação dos seguintes pacotes do R: Rcpp e inline.

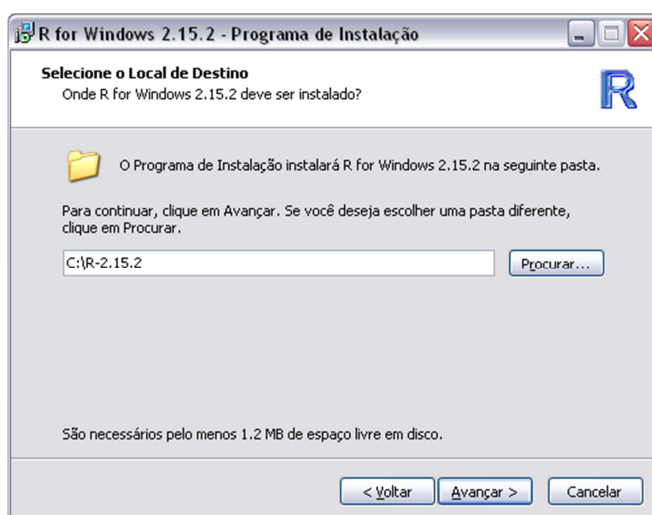
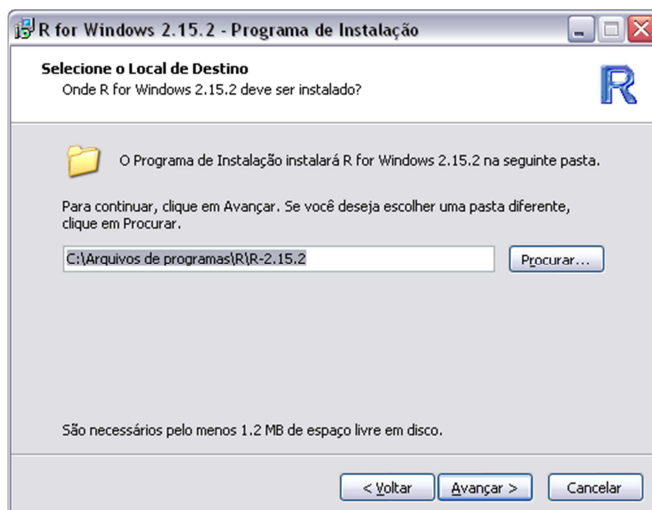
Referências:

- PINTO, Julimar Santos; CAETANO, André Junqueira. A Heterogeneidade da Vulnerabilidade Social das Juventudes: Uma Perspectiva Empírica Através do Método Grade of Membership. Revista Mediações, Londrina, v. 18, n. 1, 2013. Disponível em: <<http://www.uel.br/revistas/uel/index.php/mediacoes/article/view/16453/13217>>. Acesso em 27 set. 2013. ISSN: 1414-0543.
- CAETANO, André Junqueira; MACHADO, Carla Jorge. Consistência e identificabilidade no modelo grade of membership: uma nota metodológica. Revista Brasileira de Estudos Populacionais, Rio de Janeiro, v. 26, n. 1, p. 145-149, jan./jun., 2009.
- MANTON, Kenneth G.; WOGBURY, Max A.; TOLLEY, H. Dennis. Statistical applications using fuzzy sets. New York: John Wiley & Sons, 1994.
- WOODBURY, M. A.; CLIVE, J. Clinical pure types as a fuzzy partition. Journal of Cybernetics, 4, p. 111-121, 1974.
- WOODBURY, M. A.; CLIVE, J.; GARSON JR., A. Mathematical typology: a grade of membership technique for obtaining disease definition. Computers and Biomedical Research, 11:277-298, 1978.
- ZADEH, L. A. Fuzzy sets. Information Control. v.8, p.338-353, 1965.

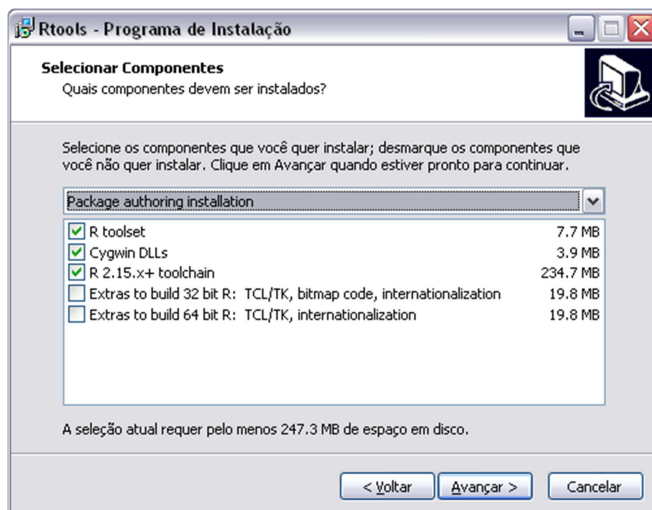
Download e Instalação do R e do Rtools:

Para a instalação do R e do Rtools é necessário realizar os downloads a partir do site oficial do projeto R: <<http://www.r-project.org/>>. Alternativamente, pode-se realizar o download diretamente do site oficial do R seguindo o link: <<http://cran.r-project.org/bin/windows/base/>>. A versão utilizada do R foi a 2.15.2 (Trick or Treat) lançada em 26/10/2012. O download do Rtools também pode ser realizado diretamente do site oficial do R: <<http://cran.r-project.org/bin/windows/Rtools/>>. Para versões do R > 2.15.1 até R 2.16.x deve-se realizar o download do Rtools216.exe.

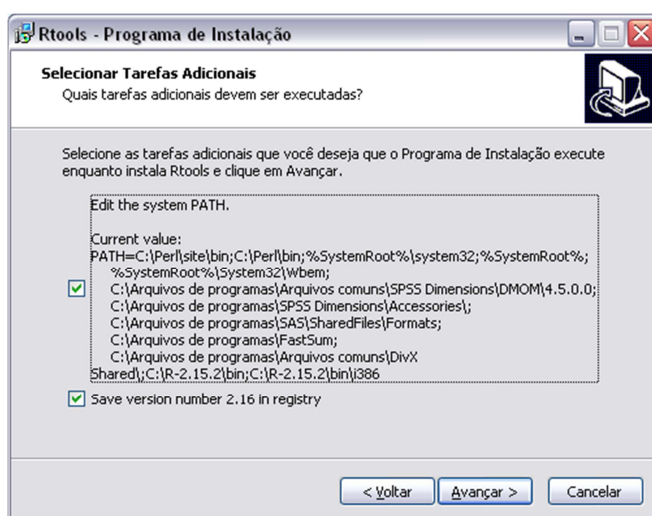
Após os downloads realize primeiro a instalação do R. Durante o processo de instalação altere o local de destino para “C:\R-2.15.2”, conforme ilustrado abaixo:




Em seguida realize a instalação do Rtools. Durante o processo de instalação do Rtools certifique-se de que pelo menos as ferramentas R toolset, Cygwin DLLs e R 2.15.x+ toolchain serão instaladas, conforme ilustrado abaixo:



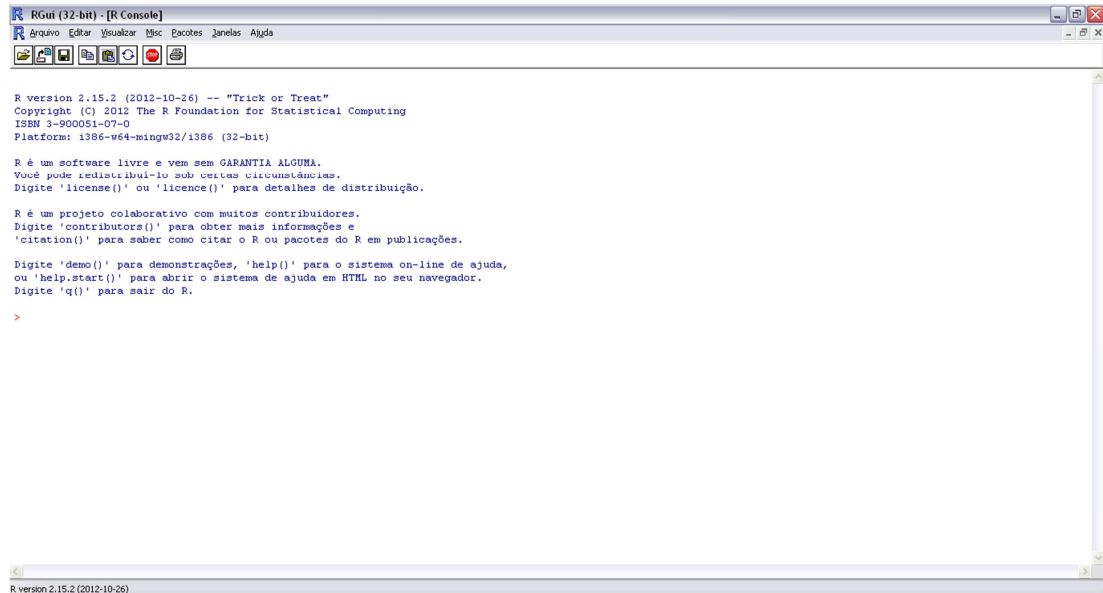
Não deixe de marcar a opção “Edit the system PATH”, conforme ilustrado abaixo:



Instalação dos pacotes requeridos

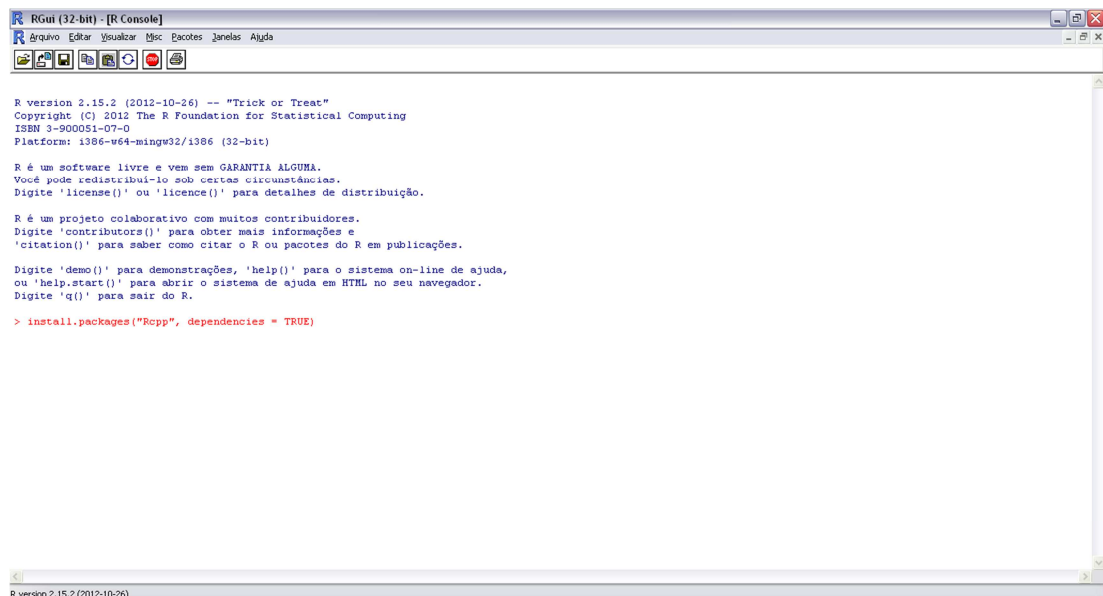
Uma vez instalado o R e o Rtools acesse o R com um duplo clique no ícone do R que está localizado na área de trabalho: 

Então, o console do R se abrirá, conforme ilustrado abaixo:



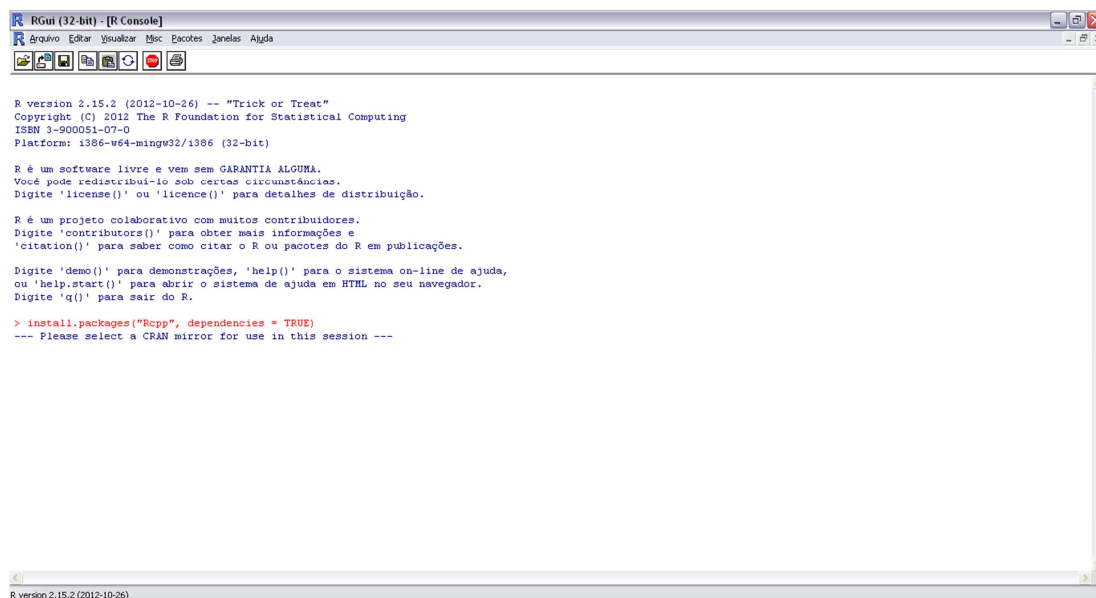
Para instalar os pacotes requeridos é necessária uma conexão de internet ativa. Copie e cole o seguinte código no console do R, conforme ilustrado abaixo:

`install.packages("Rcpp", dependencies = TRUE)`



Após colar o código tecle <ENTER>.

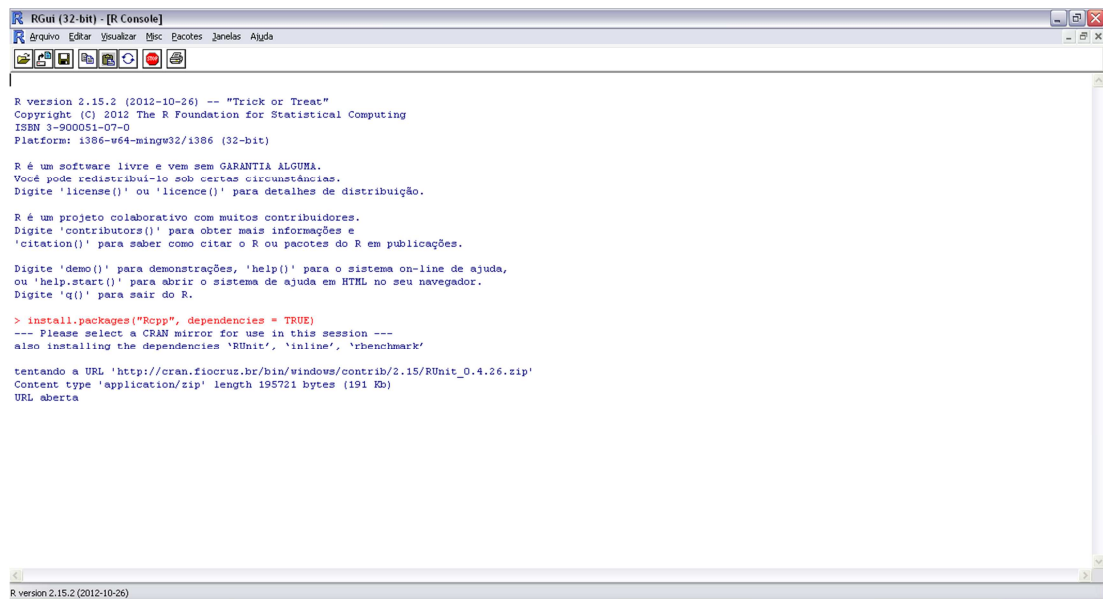
Havendo uma conexão de internet, o R lhe pedirá que selecione um CRAN, isto é, a seleção de um repositório baseado em uma rede de protocolo FTP que é mantida por servidores web, de onde as diversas versões do R, arquivos de documentação e os diversos pacotes podem ser acessados por download.



Neste caso, foi escolhido Brazil (RJ), esse é o endereço do CRAN da Fiocruz (cran.fiocruz.br). Pode-se escolher qualquer outro CRAN, lembrando apenas que servidores mais distantes proporcionam downloads mais lentos.



O R realizará primeiro os downloads das seguintes dependências: Runit, inline e rbenchmark.



```
RGui (32-bit) - [R Console]
Arquivo Editar Visualizar Plots Pacotes Janelas Ajuda

R version 2.15.2 (2012-10-26) -- "Trick or Treat"
Copyright (C) 2012 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-w64-mingw32/i386 (32-bit)

R é um software livre e vem sem GARANTIA ALGUMA.
Você pode redistribuí-lo sob certas circunstâncias.
Digite 'license()' ou 'licence()' para detalhes de distribuição.

R é um projeto colaborativo com muitos contribuidores.
Digite 'contributors()' para obter mais informações e
'citation()' para saber como citar o R ou pacotes do R em publicações.

Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.

> install.packages("Rcpp", dependencies = TRUE)
--- Please select a CRAN mirror for use in this session ---
also installing the dependencies 'RUnit', 'inline', 'rbenchmark'

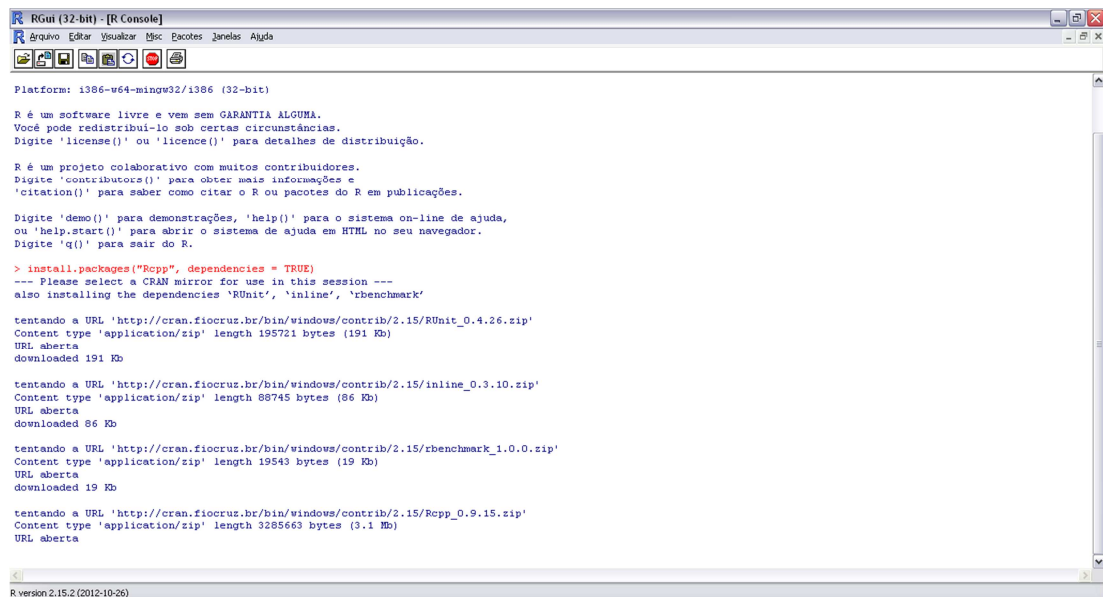
tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/RUnit_0.4.26.zip'
Content type 'application/zip' length 195721 bytes (191 Kb)
URL aberta
downloaded 191 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/inline_0.3.10.zip'
Content type 'application/zip' length 88745 bytes (86 Kb)
URL aberta
downloaded 86 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/rbenchmark_1.0.0.zip'
Content type 'application/zip' length 19543 bytes (19 Kb)
URL aberta
downloaded 19 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/Rcpp_0.9.15.zip'
Content type 'application/zip' length 3285663 bytes (3.1 Mb)
URL aberta
```

Após o download das dependências o R realizará o download do Rcpp.



```
RGui (32-bit) - [R Console]
Arquivo Editar Visualizar Plots Pacotes Janelas Ajuda

Platform: i386-w64-mingw32/i386 (32-bit)

R é um software livre e vem sem GARANTIA ALGUMA.
Você pode redistribuí-lo sob certas circunstâncias.
Digite 'license()' ou 'licence()' para detalhes de distribuição.

R é um projeto colaborativo com muitos contribuidores.
Digite 'contributors()' para obter mais informações e
'citation()' para saber como citar o R ou pacotes do R em publicações.

Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.

> install.packages("Rcpp", dependencies = TRUE)
--- Please select a CRAN mirror for use in this session ---
also installing the dependencies 'RUnit', 'inline', 'rbenchmark'

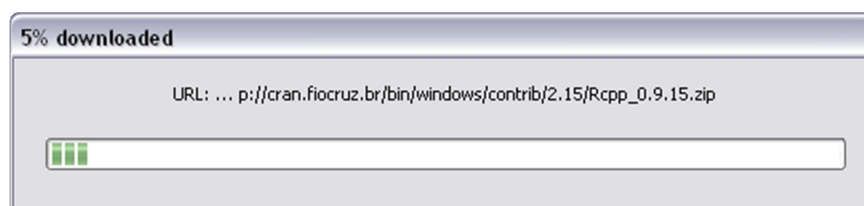
tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/RUnit_0.4.26.zip'
Content type 'application/zip' length 195721 bytes (191 Kb)
URL aberta
downloaded 191 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/inline_0.3.10.zip'
Content type 'application/zip' length 88745 bytes (86 Kb)
URL aberta
downloaded 86 Kb

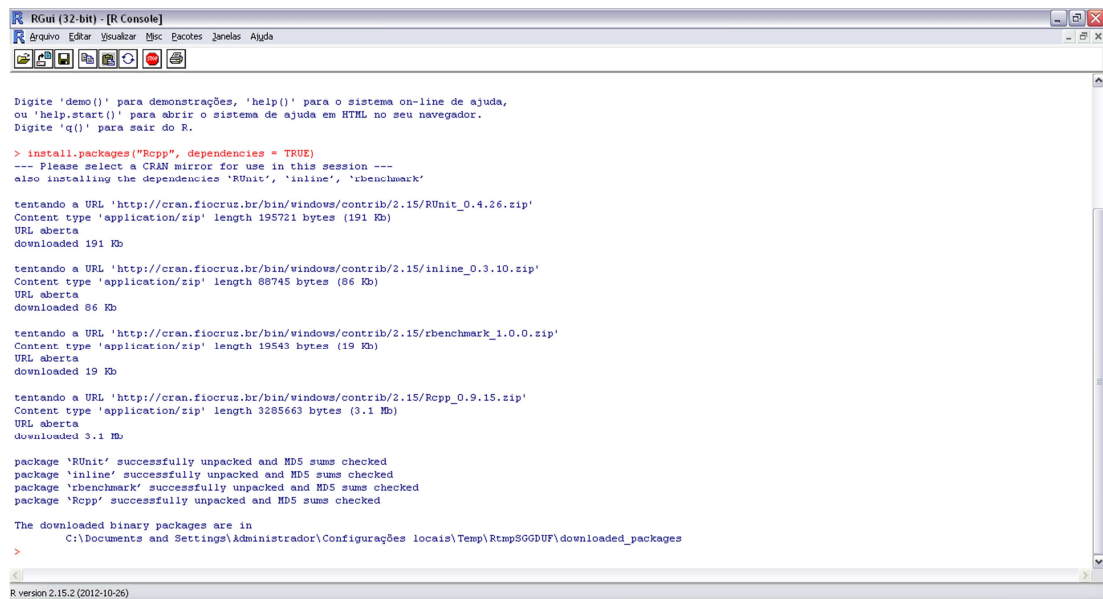
tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/rbenchmark_1.0.0.zip'
Content type 'application/zip' length 19543 bytes (19 Kb)
URL aberta
downloaded 19 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/Rcpp_0.9.15.zip'
Content type 'application/zip' length 3285663 bytes (3.1 Mb)
URL aberta
```

Para cada download o R exibe uma barra de progresso.



Após o término dos downloads dos pacotes requeridos o R descompactará cada pacote e checará a integridade de cada um utilizando o MD5 sums.



```

RGui (32-bit) - [R Console]
Arquivo Editar Visualizar Misc Pacotes Janelas Ajuda

Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.

> install.packages("Rcpp", dependencies = TRUE)
--- Please select a CRAN mirror for use in this session ---
also installing the dependencies 'RUnit', 'inline', 'rbenchmark'

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/RUnit_0.4.26.zip'
Content type 'application/zip' length 195721 bytes (191 Kb)
URL aberta
downloaded 191 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/inline_0.3.10.zip'
Content type 'application/zip' length 88745 bytes (86 Kb)
URL aberta
downloaded 86 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/rbenchmark_1.0.0.zip'
Content type 'application/zip' length 19543 bytes (19 Kb)
URL aberta
downloaded 19 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/Rcpp_0.9.15.zip'
Content type 'application/zip' length 3285663 bytes (3.1 Mb)
URL aberta
downloaded 3.1 Mb

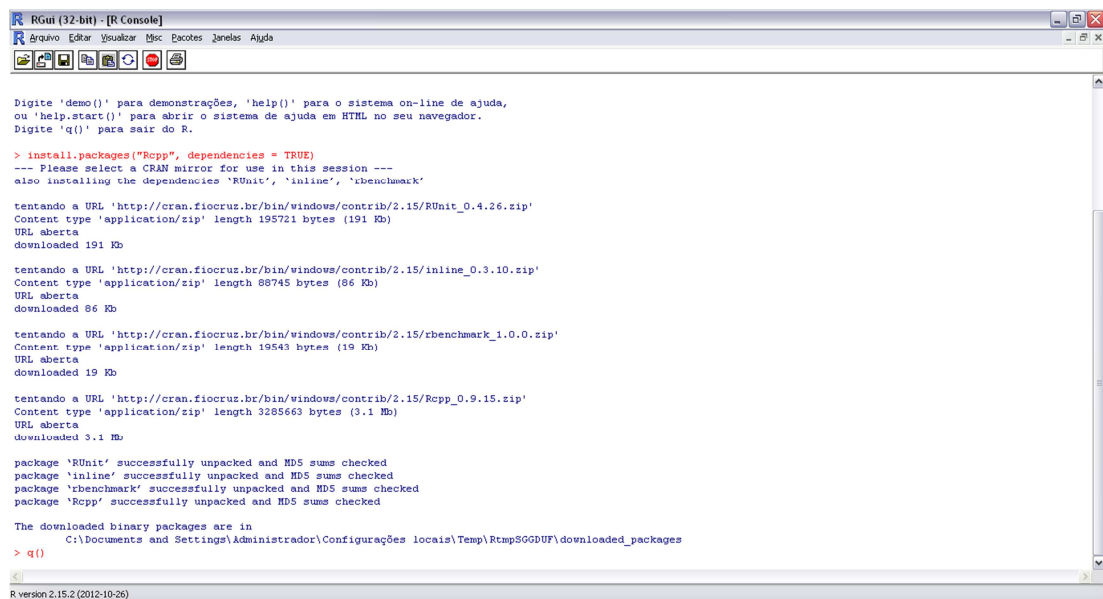
package 'RUnit' successfully unpacked and MD5 sums checked
package 'inline' successfully unpacked and MD5 sums checked
package 'rbenchmark' successfully unpacked and MD5 sums checked
package 'Rcpp' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
C:\Documents and Settings\Administrador\Configurações locais\Temp\RtmpSGGDUf\downloaded_packages

>
R version 2.15.2 (2012-10-26)

```

Terminado o processo de instalação dos pacotes, pode-se encerrar o R com o seguinte comando: q()



```

RGui (32-bit) - [R Console]
Arquivo Editar Visualizar Misc Pacotes Janelas Ajuda

Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.

> install.packages("Rcpp", dependencies = TRUE)
--- Please select a CRAN mirror for use in this session ---
also installing the dependencies 'RUnit', 'inline', 'rbenchmark'

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/RUnit_0.4.26.zip'
Content type 'application/zip' length 195721 bytes (191 Kb)
URL aberta
downloaded 191 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/inline_0.3.10.zip'
Content type 'application/zip' length 88745 bytes (86 Kb)
URL aberta
downloaded 86 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/rbenchmark_1.0.0.zip'
Content type 'application/zip' length 19543 bytes (19 Kb)
URL aberta
downloaded 19 Kb

tentando a URL 'http://cran.fiocruz.br/bin/windows/contrib/2.15/Rcpp_0.9.15.zip'
Content type 'application/zip' length 3285663 bytes (3.1 Mb)
URL aberta
downloaded 3.1 Mb

package 'RUnit' successfully unpacked and MD5 sums checked
package 'inline' successfully unpacked and MD5 sums checked
package 'rbenchmark' successfully unpacked and MD5 sums checked
package 'Rcpp' successfully unpacked and MD5 sums checked

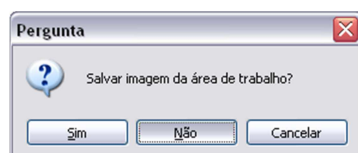
The downloaded binary packages are in
C:\Documents and Settings\Administrador\Configurações locais\Temp\RtmpSGGDUf\downloaded_packages

> q()
R version 2.15.2 (2012-10-26)

```

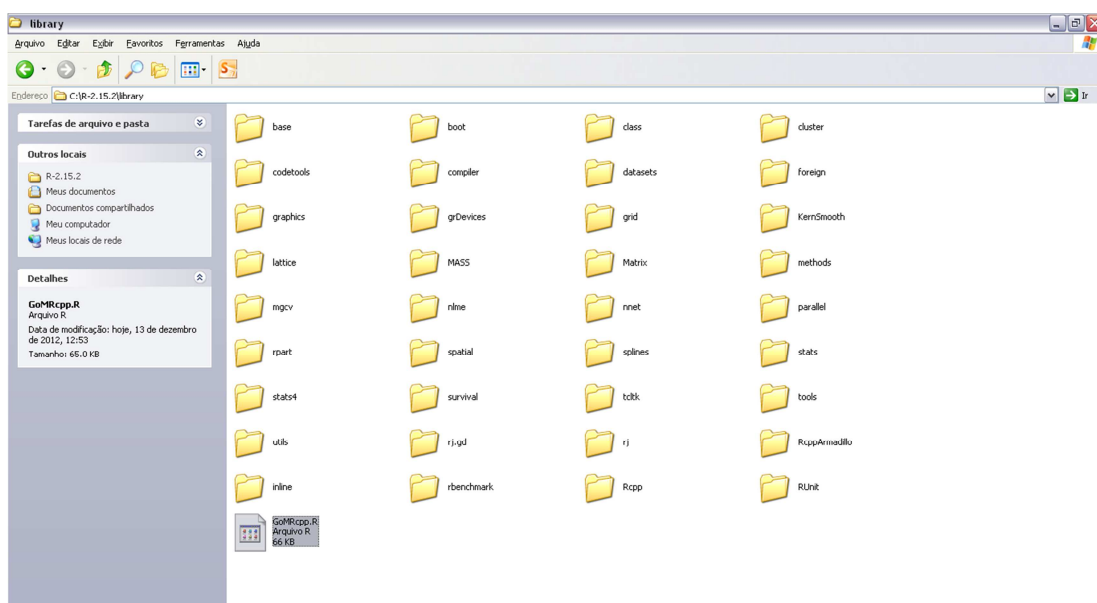
Após colar o código tecele <ENTER>.

Na janela abaixo clique em Não para encerrar o R.



Copiar GoMRcpp para a biblioteca do R

Uma vez completado os processos anteriores, ainda é necessário que o R possa localizar o arquivo GoMRcpp.R. Sugere-se que o arquivo GoMRcpp.R seja colocado dentro da pasta da biblioteca do R, isso não é uma regra. Se você houver instalado o R no caminho “C:\R-2.15.2”, então a biblioteca do R estará em “C:\R-2.15.2\library”. Mas atenção, se você houver instalado o R em outro caminho verifique se existe caractere de espaço no caminho que você instalou (e.g. “C:\Arquivos de programas\R\R-2.15.2”). Se houver caractere de espaço desinstale o R e reinstale-o novamente, não esquecendo de também reinstalar os pacotes.



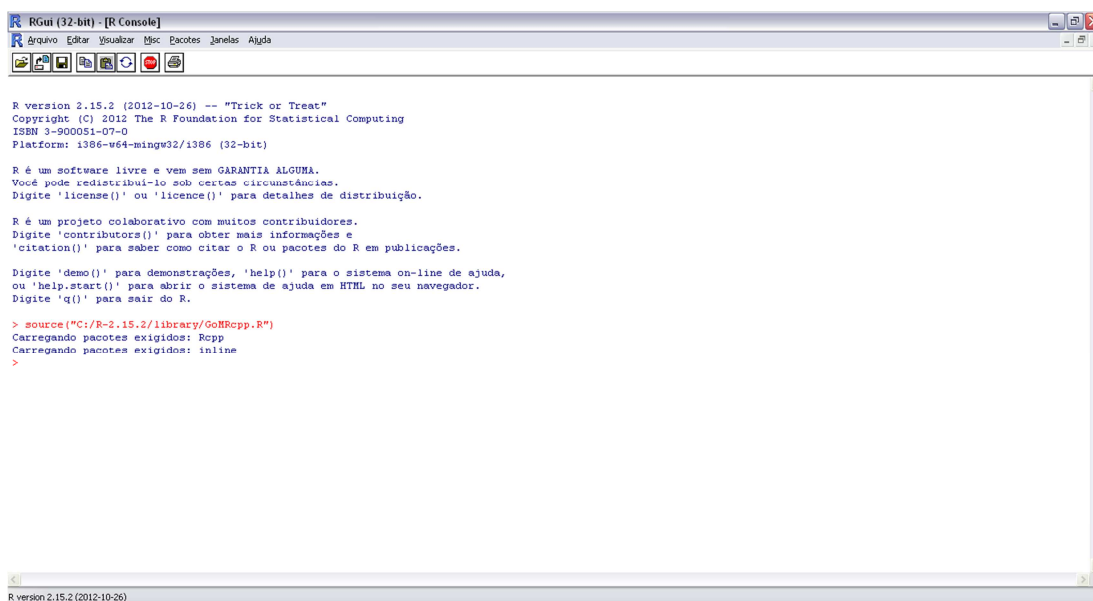
Dessa forma o caminho totalmente qualificado (ou seja, caminho + nome do arquivo) do R-Script GoMRcpp.R será “C:\R-2.15.2\library\GoMRcpp.R”, ou algo semelhante se você houver instalado o R em caminho diferente de “C:\R-2.15.2”. Assim, toda vez que se fizer referência ao R-Script GoMRcpp.R no R deve-se utilizar o caminho totalmente qualificado. Porém, o R não reconhece o caractere de barra invertida “\” como descritor de caminho de diretório ou arquivo, devendo-se utilizar dois caracteres de barra invertida (“C:\\R-2.15.2\\library\\GoMRcpp.R”) ou apenas um caractere de barra (“C:/R-2.15.2/library/GoMRcpp.R”). Se seu teclado não possui a tecla de barra “/” você pode escrevê-la utilizando a seguinte combinação de teclas: AltGr (ou seja, Alt direito) + q.

Linha de comando do R para carregar GoMRcpp

Uma vez completado os passos anteriores, com sucesso, não será mais necessário refazê-los, a menos que deseje desinstalar e reinstalar o R na mesma ou em outra versão.

Porém, TODA VEZ que desejar utilizar o R-Script GoMRcpp.R após ter iniciado uma nova seção do R faz-se necessário utilizar a linha de código abaixo para carregar no ambiente ativo do R as funções dos pacotes Rcpp e inline que, junto com o Rtools, fornecem as dependências de que precisa o R-Script GoMRcpp.R. Tanto quanto proporciona o carregamento do próprio GoMRcpp.R e de suas funções específicas para o ambiente ativo do R. Para isso a função `source()` é chamada e `lhe` é passado, como parâmetro, o caminho totalmente qualificado do R-Script GoMRcpp.R, delimitado por aspas.

```
source("C:/R-2.15.2/library/GoMRcpp.R")
```



Após colar o código tecle <ENTER>.

A função GoMRcpp

Abaixo se apresenta a função GoMRcpp() que executa o método de GoM:

```
GoMRcpp(data.object = NULL,
        initial.K = 2, final.K = initial.K,
        gamma.algorithm = c("gradient.1992", "woodbury.1974"),
        initial.gamma = c("equal.values", "random", "pure1", "gamma.object"),
        initial.gamma.object = NULL,
        gamma.fit = TRUE,
        lambda.algorithm = c("gradient.1992", "woodbury.1974"),
        initial.lambda = c("random", "pure1", "equal.values", "lambda.object"),
        initial.lambda.object = NULL,
        lambda.fit = TRUE,
        case.id = NA,
        case.weight = NA,
        internal.var = NULL,
        order.K = TRUE,
        omega.fit = FALSE,
        dec.char = ".")
```

Uma vez que o R-Script GoMRcpp.R estiver carregado no ambiente ativo do R, pode-se invocar a função GoMRcpp(). Alguns parâmetros da função GoMRcpp() são obrigatórios, por outro lado, alguns outros parâmetros são opcionais. Abaixo se apresenta cada um dos parâmetros dessa função, indicando quais são obrigatórios e quais não são.

data.object Este parâmetro da função GoMRcpp() é obrigatório e indica um objeto do tipo data.frame que deverá conter a matriz dos dados que serão modelados.

initial.K Este parâmetro é opcional e indica o número inicial de perfis que será modelado. O valor mínimo para esse parâmetro é igual a 2 (dois) perfis. Se este parâmetro não for informado a função GoMRcpp() irá especificar por padrão o valor 2 (dois), que é o número mínimo de perfis que podem ser estimados pelo método de GoM.

final.K Este parâmetro é opcional e indica o número final de perfis que será modelado. O valor mínimo para esse parâmetro é igual a 2 (dois) perfis. Se este parâmetro não for informado a função GoMRcpp() irá especificar por padrão o mesmo valor informado para o parâmetro initial.K, portanto executará apenas uma modelagem de GoM para um valor fixo de K.

gamma.algorithm Este parâmetro é opcional e indica o algoritmo que será utilizado para a estimação dos gamas (g_{ik}). Se este parâmetro não for informado a função `GoMRcpp()` irá especificar por padrão o algoritmo do gradiente (“gradient.1992”). A outra opção deste parâmetro (“woodbury.1974”) emprega o algoritmo originalmente construído para o GoM.

initial.gamma Este parâmetro é opcional e indica a forma de composição da matriz inicial que será utilizada para a estimação dos gamas (g_{ik}). Se este parâmetro não for informado a função `GoMRcpp()` irá especificar por padrão a forma de valores iguais para a composição da matriz inicial. As outras opções são: i) a forma de composição da matriz inicial será gerada aleatoriamente (“random”); ii) a forma de composição da matriz inicial será gerada de modo que os indivíduos possuam pertencimento pleno, i.e. g_{ik} igual a 1, para o perfil $k=1$ (“pure1”) e, portanto, g_{ik} igual a 0 (zero) para os demais valores de k ; iii) a forma de composição da matriz inicial será atribuída pelo usuário (“gamma.object”).

initial.gamma.object Este parâmetro é opcional. Porém, caso o usuário informe que o valor do parâmetro `initial.gamma` seja igual a “gamma.object”, então será obrigatório indicar para esse parâmetro um objeto do tipo `data.frame` que deverá conter a matriz inicial dos dados dos gamas (g_{ik}) que serão modelados.

gamma.fit Este parâmetro é opcional e indica um valor booleano (TRUE/FALSE). Se este parâmetro não for informado a função `GoMRcpp()` irá especificar por padrão o valor TRUE, indicando que os gamas serão modelados pelo processo iterativo. Se FALSE os valores dos gamas iniciais não sofrerão alteração, já que não serão modelados pelo processo iterativo.

lambda.algorithm Este parâmetro é opcional e indica o algoritmo que será utilizado para a estimação dos lambdas (λ_{kjl}). Se este parâmetro não for informado a função `GoMRcpp()` irá especificar por padrão o algoritmo do gradiente (“gradient.1992”). A outra opção deste parâmetro (“woodbury.1974”) emprega o algoritmo originalmente construído para o GoM.

initial.lambda Este parâmetro é opcional e indica a forma de composição da matriz inicial que será utilizada para a estimação dos lambdas (λ_{kjl}). Se este parâmetro não for informado a função `GoMRcpp()` irá especificar por padrão a forma de valores iguais para a composição da matriz inicial. As outras opções são: i) a forma de composição da matriz inicial será gerada aleatoriamente (“random”); ii) a

forma de composição da matriz inicial será gerada de modo que as categorias (I) de menor valor para cada variável (j) possuam probabilidades iguais a 1, para o perfil $k=1$ (“pure1”) e, portanto, λ_{kjl} será igual a 0 (zero) para as demais categorias de cada variável j do perfil $k=1$. Para os demais perfis, ou seja, perfis $k>1$ os valores de λ_{kjl} possuirão valores iguais; iii) a forma de composição da matriz inicial será atribuída pelo usuário (“lambda.object”).

initial.lambda.object Este parâmetro é opcional. Porém, caso o usuário informe que o valor do parâmetro initial.lambda seja igual a “lambda.object”, então será obrigatório indicar para esse parâmetro um objeto do tipo array que deverá conter a matriz inicial dos dados dos lambdas (λ_{kjl}) que serão modelados.

lambda.fit Este parâmetro é opcional e indica um valor booleano (TRUE/FALSE). Se este parâmetro não for informado a função GoMRcpp() irá especificar por padrão o valor TRUE, indicando que os lambdas serão modelados pelo processo iterativo. Se FALSE os valores dos lambdas iniciais não sofrerão alteração, já que não serão modelados pelo processo iterativo.

case.id Este parâmetro é obrigatório e indica o nome de uma variável identificação de cada elemento da matriz de dados.

case.weight Este parâmetro é opcional e indica o nome de uma variável de peso que pondera cada elemento da matriz de dados.

internal.var Este parâmetro é opcional e indica o nome das variáveis que serão consideradas na modelagem do método de GoM. Se este parâmetro não for informado a função GoMRcpp() irá especificar por padrão um vetor que conterá todos os nomes de variáveis contidas na matriz de dados, com exceção do nome da variável informada no parâmetro case.id e do nome da variável informada no parâmetro case.weight, caso esta última seja informada.

order.K Este parâmetro é opcional e indica um valor booleano (TRUE/FALSE). Se este parâmetro não for informado a função GoMRcpp() irá especificar por padrão o valor TRUE, indicando que será realizada uma tentativa de organizar os perfis segundo os valores apresentados por λ_{kjl} nas categorias (I) das variáveis (j).

omega.fit Este parâmetro é opcional e indica um valor booleano (TRUE/FALSE). Se este parâmetro não for informado a função GoMRcpp() irá especificar por padrão o valor FALSE. Ao especificar o valor TRUE será criada uma

matriz que conterá todos os possíveis padrões de resposta ao longo de todas as categorias (l) para todas as variáveis (j). Essa matriz será mesclada com os dados da matriz informada no parâmetro `data.object`. A possibilidade de uso desse parâmetro está condicional à quantidade máxima de registros que o ambiente R pode processar, dado por 2^{31} .

dec.char Este parâmetro é opcional e indica o caractere decimal que será utilizado nos arquivos de saída: `GoMKi(x).TXT`, `LogGoMKi(x).TXT`. Onde, neste caso, o subscrito _i indica a quantidade de perfis que foram modelados pelo método de GoM e o subscrito _x indica a ordem de execução do modelo de GoM de _i perfis. Se este parâmetro não for informado, a função `GoMRcpp()` irá especificar por padrão o valor “.” (ponto).

Este arquivo passo-a-passo acompanha o R-Script `GoMRcpp.R` na versão 01.01.00 que possui o seguinte cabeçalho:

```
#####
## Title: GoMRcpp                                                    ##
## Version: 01.01.00                                                ##
## Date: 2013-31-07                                                 ##
## Autor: Julimar Pinto, Andre Caetano                             ##
## Maintainer: Julimar Pinto <julimarsp.r@gmail.com>                ##
## Description: GoMRcpp.R is a R-Script file for modeling Grade of Membership (GoM) to discrete data ##
## Depends: R (>= 2.15.2); *Rtools (>= 2.16.0.1923); **gcc; Rcpp (>= 0.9.15); inline (>= 0.3.10) ##
## *Only for Windows Operating System                               ##
## **Only for Linux Operating System                                ##
#####
```