

R のインストールと超入門

Rとは？

Rダウンロード

Rのインストール

Rの基本操作

R終了

データの読み込み

パッケージの操作

Rとは？

Rはデータ解析・マイニングを行うフリーソフトである。Rはデータ解析の環境でもあり、言語でもある。ニュージーランドのオークランド(Auckland)大学の統計学科のRobert GentlemanとRoss Ihakaにより開発がはじめられ、1997年からは多くの賛同者が加えられ、開発され続けているオープンソース方式のデータ解析・処理の専用ソフトである。

R本体は、データの簡単な操作環境、プログラミング環境、パッケージの操作環境を提供する。

複雑のデータ解析やマイニングを行うためには、既存のパッケージをインストールして用いるか、それともプログラミング組むことになる。

Rは平均年2回のバージョンアップを行っている。2009年3月の現在の最新バージョンは2.8.1である。

Rに用いる数千個のパッケージがインターネット上で公開されている。Rのユーザは、近年急速に増えている。

Rの大きいメリットは、フリーであることと、データ解析のパッケージの豊富さである。

欠点は、コマンドで主な操作を行う点である。これは、ExcelやSPSSのようにメニューとシートの操作に慣れている方に対しては慣れるまでは苦痛に感じるであろう。しかし、脳のトレーニングとしては、RはExcel、SPSSのようなデータ解析ソフトよりよいであろう。

Rのダウンロード

Rの生まれた故郷は、ニュージーランドであるが、多くの国にミラーサイトがある。R本家のサイトは <http://www.r-project.org/> である。

日本では、会津大学、東京大学、筑波大学にミラーサイトがある。

<ftp://ftp.u-aizu.ac.jp/pub/lang/R/CRAN/bin/>

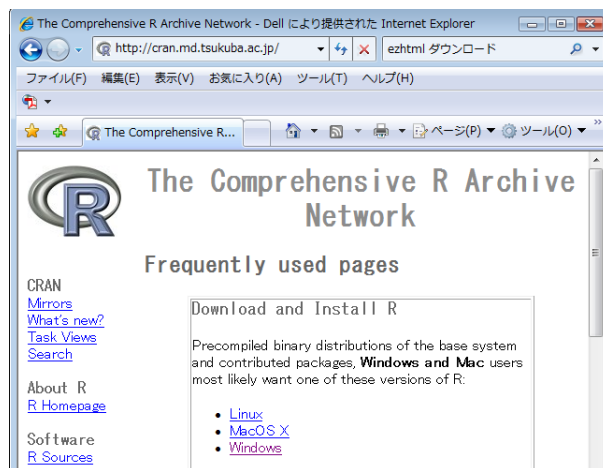
<ftp://ftp.ecc.u-tokyo.ac.jp/CRAN/>

<http://cran.md.tsukuba.ac.jp/>

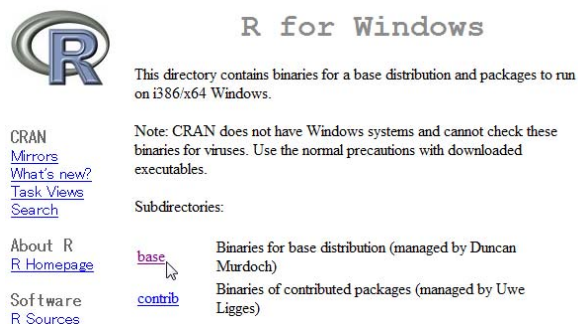
R はこれらのサイトからダウンロードすることができる。R は、UNIX, Windows, Mac など主な OS に対応するバージョンがある。

ここでは、筑波大学のミラーサイトから Windows 32 ビットマシン用をダウンロードしてインストールする手順を説明する。

(1) Webブラウザを起動し、<http://cran.md.tsukuba.ac.jp/>にアクセスすると次画面が開かれる。



(2) [Windows]をクリックすると次の画面が開かれる。



(3) [base]をクリックすると次の画面が開かれ、最新バージョンのダウンロードページに辿りつく。



R-2.8.1 for Windows

CRAN
[Mirrors](#)
[What's new?](#)
[Task Views](#)
[Search](#)

[About R](#)
[R Homepage](#)

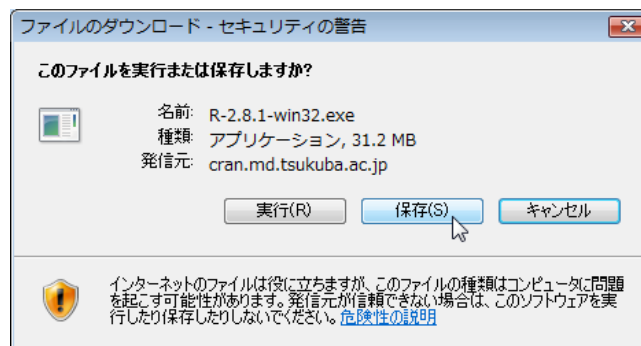
[Download R 2.8.1 for Windows](#) (34 megabytes)

[Installation and other instructions](#)

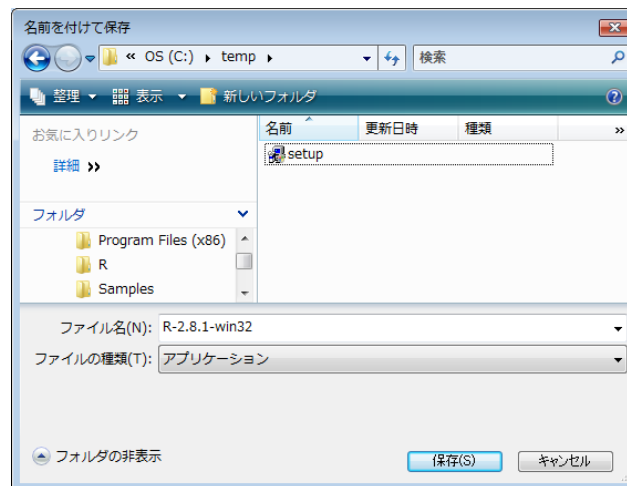
New features in this version: [Windows specific](#), [all platforms](#).

If you want to double-check that the package you have downloaded exactly matches the package distributed by R, you can compare the [md5sum](#) of the .exe to the [true fingerprint](#). You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

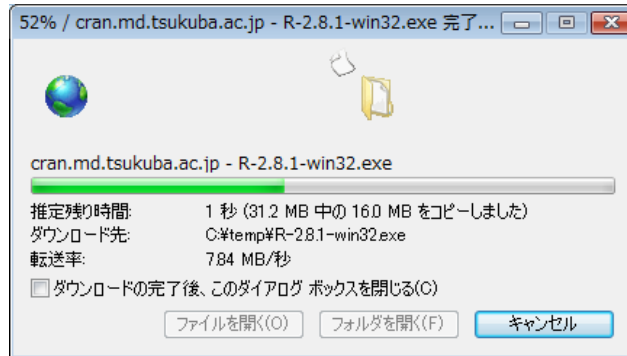
- (4) [Download R2.*.* for Windows]をクリックするとダウンロードのダイアログボックスが開かれる。



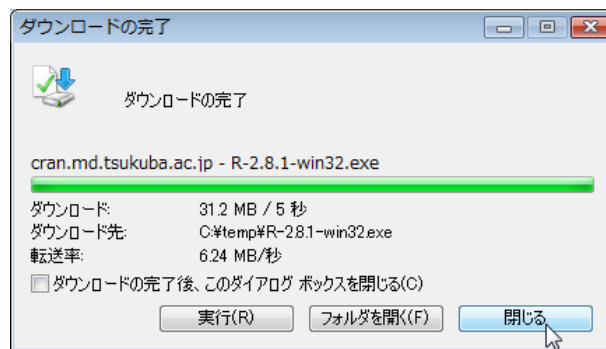
- (5) [保存(s)] ボタンを押すと保存場所を指定するダイアログボックスが開かれる。



- (6) 保存場所を指定し、画面の[保存(s)]のボタンを押すとダウンロードが始まる。



- (7) ダウンロードが終わると「ダウンロードの完了」画面が開かれる。[閉じる]ボタンを押すとダウンロード作業が終了する。

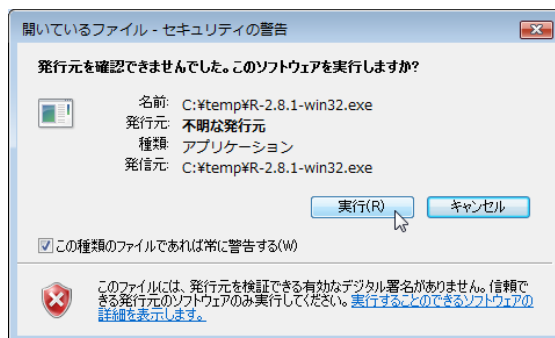


Rのインストール

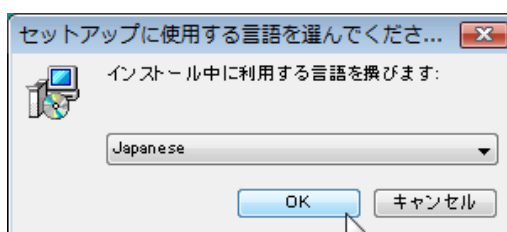
- (1) ダウンロードして保存したフォルダにアクセスするとダウンロードしたセットアップファイル [R-2.*.*-Win32]が確認できる。



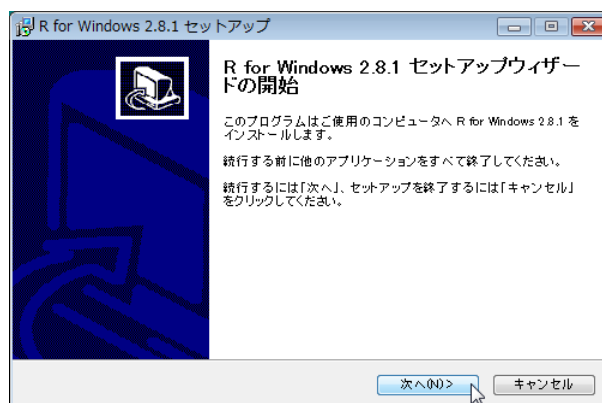
(2) セットアップファイル[R-2.*.*-Win32]をクリックするとインストールの最初画面が開かれる。



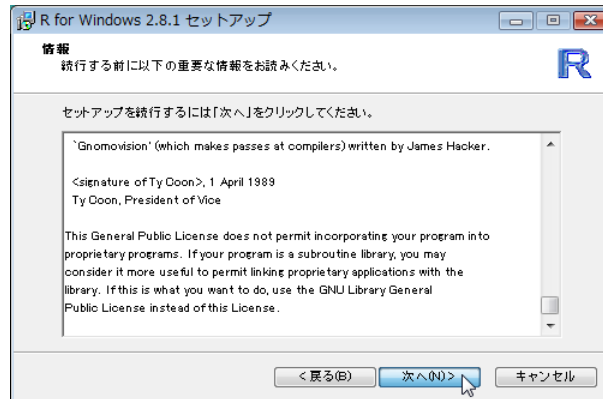
(3) インストールの最初画面の[実行(R)]ボタンを押すと次の画面が開かれる。



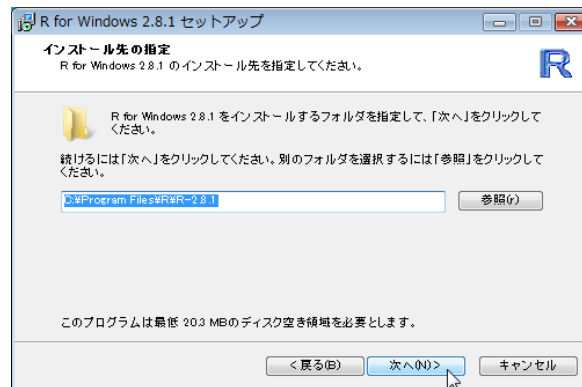
(4) 日本語が指定されているので、[OK]ボタンを押すと次の画面が開かれる。



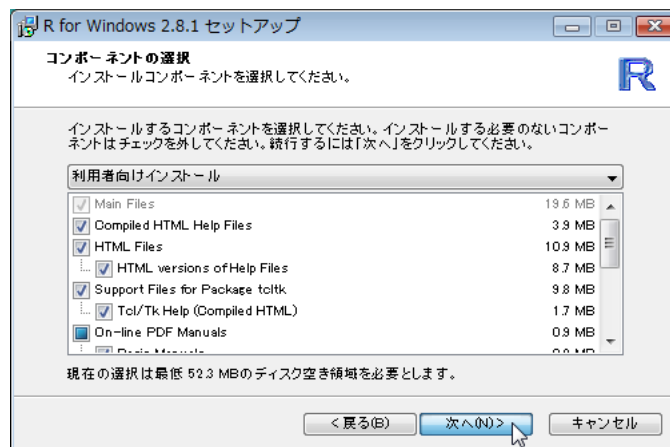
(5) 指示の通り他のアプリケーションを終了し、[次へ(N)>]ボタンを押すと次の画面が開かれる。



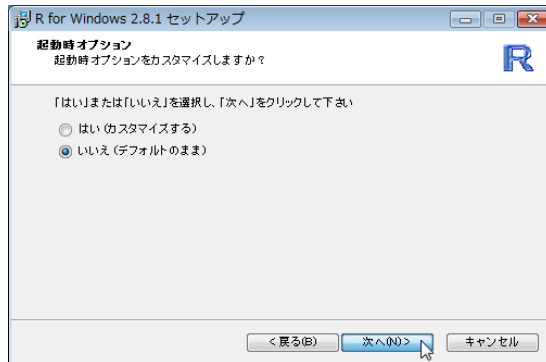
(6) [次へ(N)>]ボタンを押すと次の画面が開かれる。



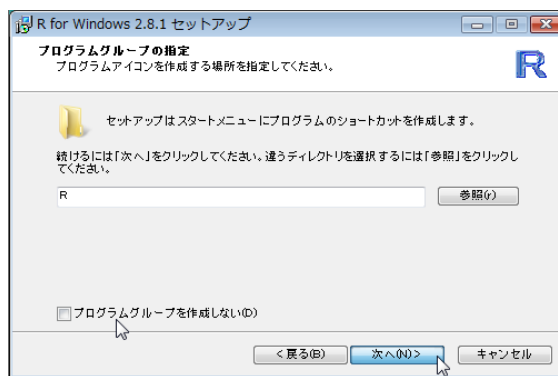
(7) インストールするフォルダを指定し、[次へ(N)>]ボタンを押すと次の画面が開かれる。



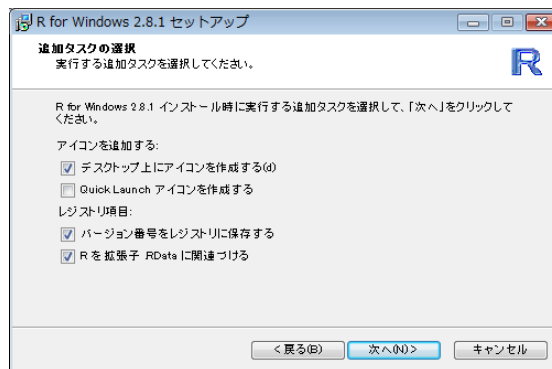
(8) インストールする項目をチェックし、[次へ(N)>]ボタンを押す。初心者は特にしてせずに、そのまま[次へ(N)>]ボタンを押すことを薦める。



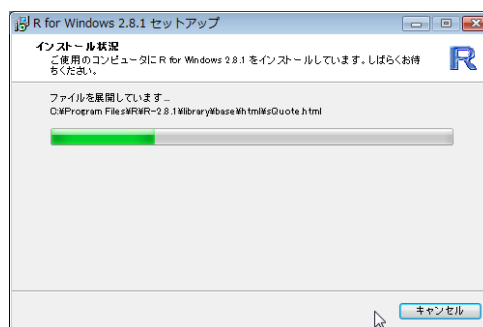
(9) そのまま[次へ(N)>]ボタンを押すと次の画面が開かれる。



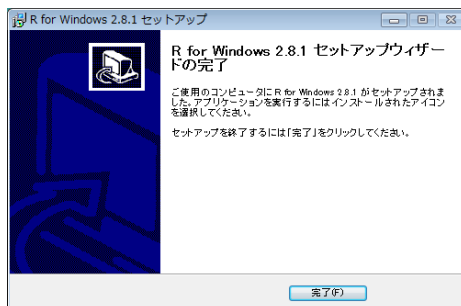
(10) そのまま[次へ(N)>]ボタンを押すと次の画面が開かれる。



(11) そのまま[次へ(N)>]ボタンを押すとインストールが始まる。



- (12) インストールが終了すると次のお画面が開かれる。




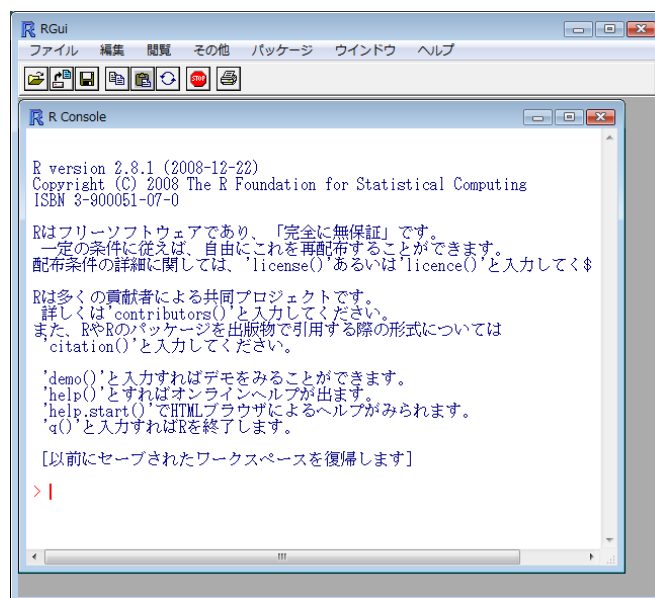
- (13) [完了(F)]ボタンを押して終了する。インストールが成功するとデスクトップにRのアイコンのコピーが作成される。



Rの基本操作

Rの起動

デスクトップのアイコンをクリックするとRが開かれる。開かれた初期画面を次に示す。



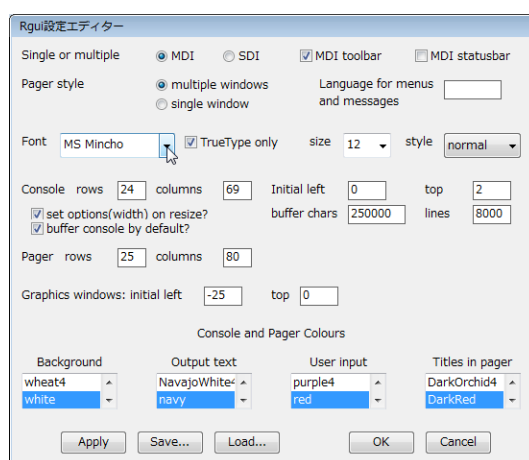
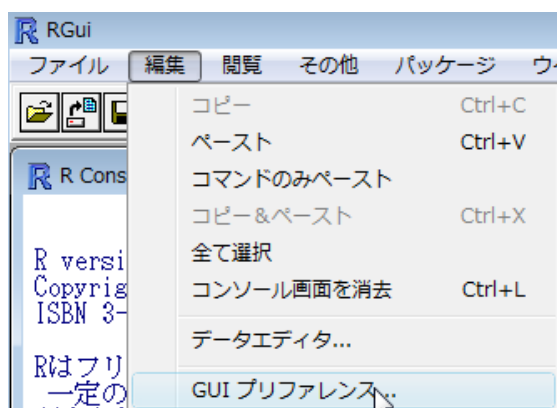
Rは画面の上部の「メニュー」と赤色の「>」の右側に数値やデータ処理の関数をキーボードで入力する方法でデータの作成や処理を行う。赤色の「>」がある画面をRのコンソール画面と呼ぶ。略してコンソールと呼ぶ。

開かれた画面の日本語文字が文字化けしたときは、まず画面の上部のメニューの「編集」→「GUI フリファレンス...」をクリックし、Rgui の設定エディターを開く。

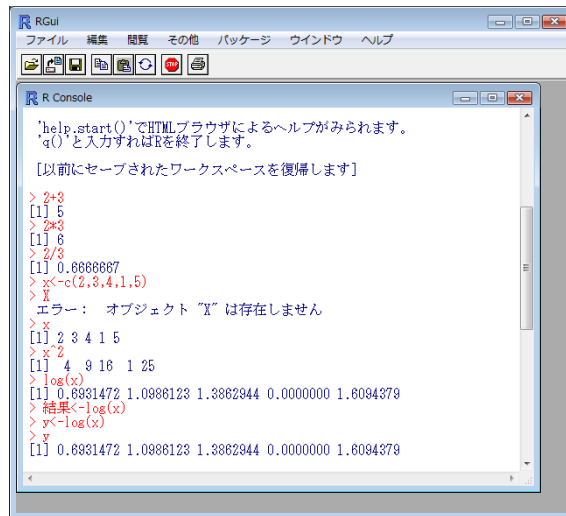
次に、Rgui の設定エディター画面で文字フォントを日本語の文字フォント MS Mincho か MS Gothic を指定する。

続いて、画面下部の[Save...]ボタンを押して保存する。保存するフォルダは特に指定せずに開かれたフォルダで構わない。


保存が終わったら、[OK]ボタンを押し、Rgui の設定エディターを終了する。



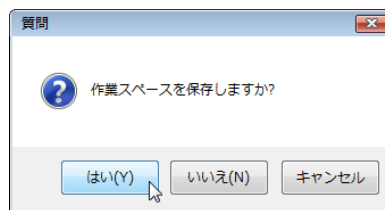
Rのコンソールデータの入力と計算処理を行う。その画面コピーを次に示す。赤色はキーボードから入力したものであり、ブル色は返された結果である。



Rの終了

メニューの「ファイル」→「終了」をクリックするか、RGui 画面の右上の  をクリックすると、作業スペースの保存を確認する画面が開かれる。

保存しないと作業内容が保存されない。[はい(Y)]のボタンを押すと作業内容を保存し、Rが終了する。



データの読み込み

小規模のデータであればコンソール上で直接入力することができる。

パッケージの操作

R は、R 上に作動するパッケージの使用環境を提供している。2009年4月の現在 CRAN というサイトだけで1700以上のパッケージが公開されている。CRAN 以外に、生物情報学(bioinformatics)関連のパッケージが [BioConductor](http://BioConductor.org) というサイトに多数公開されている。正確な統計はとっていないが、流通されているパッケージは2000を超えているのは確かである。

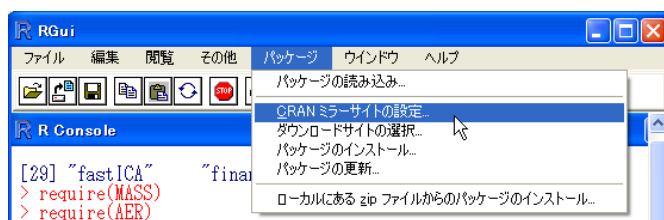
R をインストールする際に、広く使用されているいくつかのパッケージは自動的にインストールされている。これらに関しては、読み込む操作さえすれば、パッケージを利用することが可能である。インストールされているパッケージは次のコマンドで確認することができる

> installed.packages()

パッケージの読み込み(ロード)作業は、関数 `library`、`require` を用いる。パッケージ MASS を読み込む例を次に示す。

> library(MASS)

パッケージのインストールとロードはメニューバーの「パッケージ」を用いて行うことが可能である。



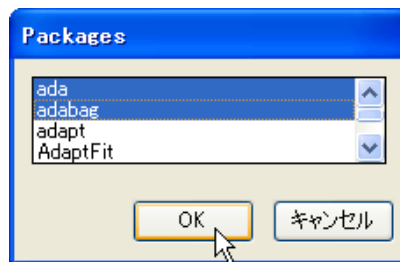
パッケージのインストールは、パッケージをローカルにダウンロードしてインストール方法とインターネット上で公開しているサイトから直接ネット上でインストールする方法がある。

ローカルにダウンロードしたパッケージをインストールするには、上記のメニューの「ローカルにある zip ファイルからのパッケージのインストール」をクリックし、zip 形式で圧縮されているパッケージを読み込むことでインストールできる。

インターネット上でインストールするためには、まずダウンロードサイトを選択することが必要である。パッケージの数が最も多いのは CRAN である。多くの国々にミラーサイトを置いてある。メニュー「パッケージ」⇒「CRAN ミラーサイトの設定」をクリックすると次のミラーサイトのリスト画面が開かれる。



現時点では、日本には3つのミラーサイトがある。上記のように一つを選択して [OK] ボタンを押し、メニュー「パッケージ」⇒「パッケージのインストール」をクリックすると次のようなパッケージのリスト画面が開かれる。



一つあるは複数のパッケージを選択し、[OK] ボタンを押すとインストールが始まる。

大学や企業団体のネットワークがインターネット接続にプロキシ(proxy)を用いる場合は、次のような1行のコマンドを実行する必要がある。たとえば、某職場のプロキシの IP アドレスが proxy.syokuba.ac.jp でポート番号が 2020 である場合は、次のようなコマンドを実行する必要がある。

```
> Sys.setenv("http_proxy"="http:// proxy.syokuba.ac.jp:2020 ")
```

コンソール上で、コマンドを用いてパッケージをインストールすることも可能である。例として、パッケージ randomForest と ada4 をインストールするコマンドを次に示す。

```
> install.packages(c("randomForest","ada4"))
```

筑波大学のミラーサイト <http://cran.md.tsukuba.ac.jp/> からパッケージをインストールするコマンドの例を次に示す。

>

```
install.packages(c("randomForest","ada4"),repos="http://cran.md.  
tsukuba.ac.jp/")
```

パッケージのアンインストール関数 `remove.packages` を用いる。