MBatch 01 Installing on OSX
Tod Casasent and Andre Schultz
2018-06-22-1600

1 Introduction

These instructions are aimed at people very familiar with OS X and R administration, with administrator access to their environment. If using the versions discussed here, it may be possible to install MBatch with minimal administration knowledge.

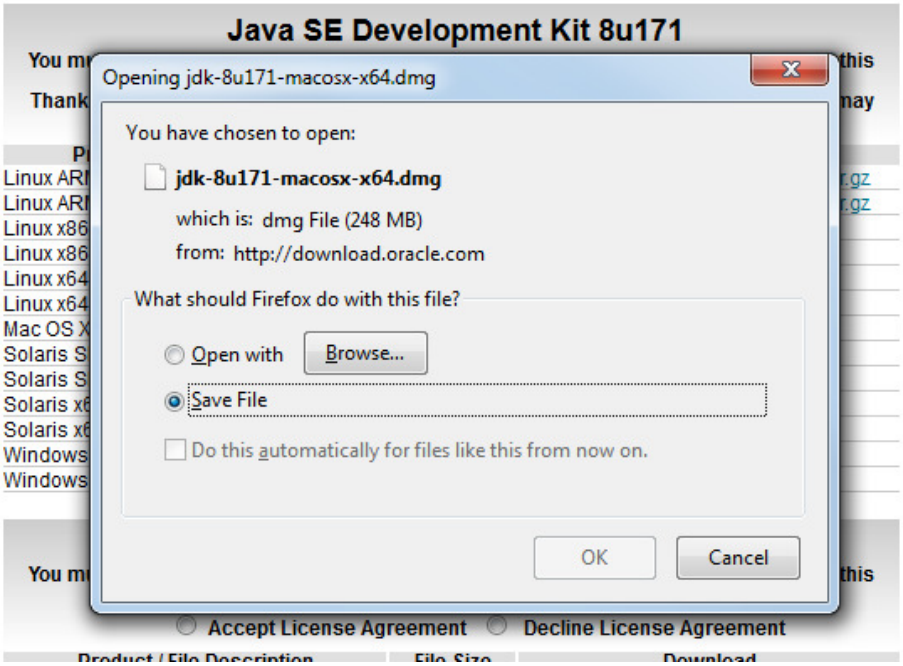
2 Target Operating System and Installation

These instructions were tested on OSX 10.9.x. These instructions with appropriate modifications should work as a basis for installing MBatch on other systems.

Perform the installs in the order given in this document. Some steps are dependent on previous steps.

3 JDK

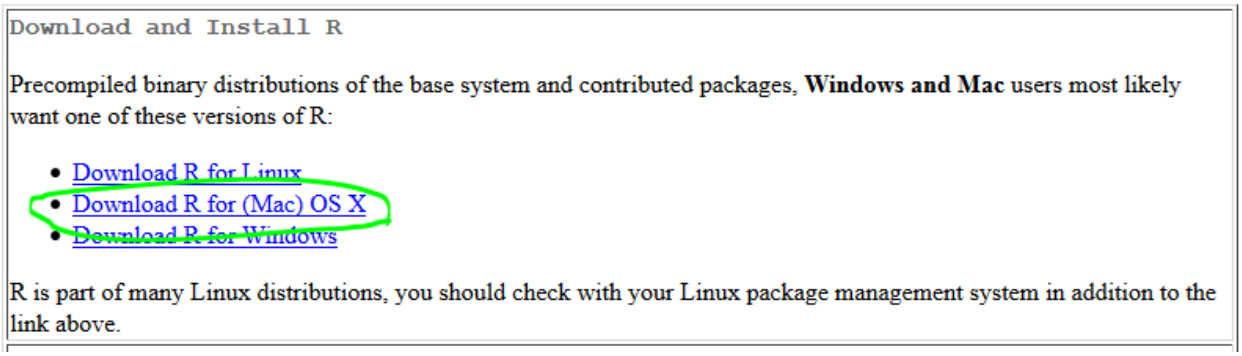
For OS X, we use the OS X x64 installation instructions and installer from http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html. Download and install the JDK8 as per Oracle's instructions.



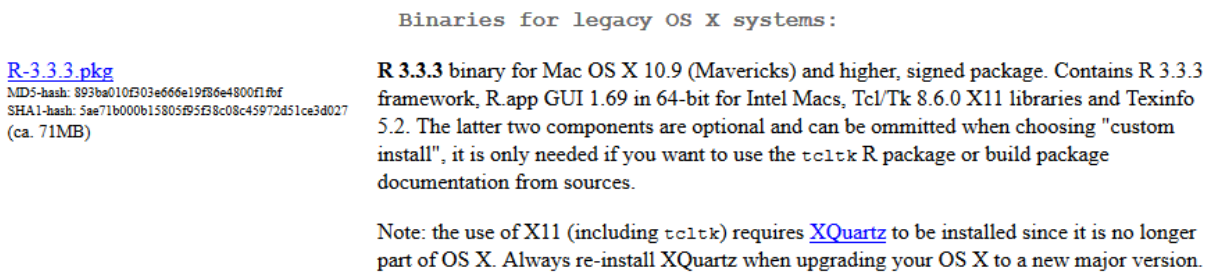
Follow the installation instructions from Oracle.

4 Install R 3.3.x

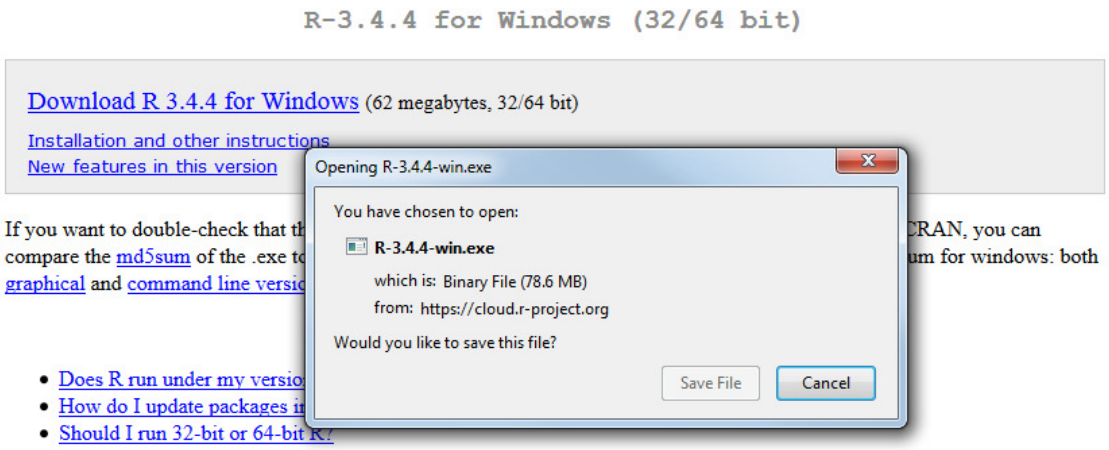
MBatch is tested and used with R 3.4.x on other systems. For OSX, we tested with 10.9.x, which limits us to installing R 3.3.3. It may work with older or newer versions, but packages and R itself tend to change between releases. Use the instructions on a local CRAN mirror (https://cran.r-project.org/mirrors.html) under "Download R for (Mac) OS X" (as shown in the screenshot).



Then use the “R-3.3.3.pkg” link under "Binaries for legacy OS X systems".



Download the R install program and follow the instructions provided under “Installation and other instructions”.



Follow the installation instructions from R-Project.

5 Setting Up a Personal R Library

If you have not installed any packages in R, you will need to create a personal package for R libraries. Install a package, as shown below, and create a personal library directory. This will appear similar to the below.

```
> install.packages("devtools", repos = "http://cran.revolutionanalytics.com")
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)
Warning in install.packages("devtools") :
  'lib = "/usr/local/lib/R/site-library"' is not writable
Would you like to use a personal library instead?  (y/n) y
Would you like to create a personal library
~/R/x86_64-pc-linux-gnu-library/3.4
to install packages into?  (y/n) y
trying URL 'https://cran.revolutionanalytics.com/src/contrib/devtools_1.13.3.tar.gz'
Content type 'application/octet-stream' length 486402 bytes (475 KB)
==================================================
downloaded 475 KB
```

6 Installing Required Packages

MBatch has some required packages.

6.1 rJava Setup for R

For OS X, rJava is installed from the R command line using the command shown below. The other lines demonstrate a test to make sure the install finished properly.

```
# install rJava from CRAN (this installs to site-library so it is available for all users)
# update to preferred CRAN mirror if desired
# this is at the R command
install.packages("rJava", dependencies=TRUE, type="source", repos = "http://cran.revolutionanalytics.com")

# check that the Java version returns properly
# should print something like: [1] "1.8.0_144"
# this is at the R command
library(rJava)
.jinit()
J("java.lang.System")$getProperty("java.version")
```

6.2 Other Required Packages

Within R, run the following installs.

```
# install packages from CRAN (this installs to site-library so it is available for all users)
# update to preferred CRAN mirror if desired
# this is at the R command
install.packages("Cairo", dependencies=TRUE, type="source", repos = "http://cran.revolutionanalytics.com")
install.packages("epiR", dependencies=TRUE, type="source", repos = "http://cran.revolutionanalytics.com")
install.packages("gtools", dependencies=TRUE, type="source", repos = "http://cran.revolutionanalytics.com")
install.packages("mclust", dependencies=TRUE, type="source", repos = "http://cran.revolutionanalytics.com")
install.packages("squash", dependencies=TRUE, type="source", repos = "http://cran.revolutionanalytics.com")
install.packages("httr", dependencies=TRUE, type="source", repos = "http://cran.revolutionanalytics.com")

# install packages from Bioconductor (this installs to site-library so it is available for all users)
# if asked about updating a/s/n, choose "a" (all)
# this is at the R command
source("http://bioconductor.org/biocLite.R")
biocLite("limma", ask="a")
biocLite("RBGL", ask="a")
biocLite("graph", ask="a")
biocLite("Biobase", ask="a")
install.packages(c("oompaBase", "ClassDiscovery", "PreProcess"), dependencies=TRUE,
repos=c("http://cran.revolutionanalytics.com", "http://silicovore.com/OOMPA/"))

# install Vennerable from github (this installs to site-library so it is available for all users)
# update to preferred CRAN mirror if desired
# this is at the R command
install.packages("devtools", repos = "http://cran.revolutionanalytics.com")
library(devtools)
install_github("js229/Vennerable")

#Check that all packages installed properly. Should return [1] TRUE
all(c("Cairo","epiR","gtools","mclust","squash","httr","oompaBase","ClassDiscovery","PreProcess","devtools","lim
ma","RBGL","graph","Biobase") %in% installed.packages()[,1])
```

7 Install MBatch Package

MBatch can be installed from GitHub using the following commands. (The devtools package was installed earlier.) Do not use the install_git option from the Linux instructions.

```
# Within R
library(devtools)
devtools::install_github("MD-Anderson-Bioinformatics/MBatch/package")
```

Note: As a side note, the "tests" directory installed via devtools is not by default included in install.packages or R CMD INSTALL. If installing MBatch via a different method, to include tests via the command line, use "--install-tests".

7.1 Populating Data for Tests

The zip file located at http://tcgadata.mdanderson.org/std_archives/MATRIX_DATA.zip contains the data used by the tests. Download the ZIP.

The zip file located at zip file located at http://tcgadata.mdanderson.org/std_archives/COMPARE.zip contains the data used to confirm the tests. Download the ZIP.

In your downloads directory, create a directory names "output".

7.2 Running the Tests

Finally, run this code to perform the tests. This takes an hour or so to run. Machines with slower processors, slower hard drives, or minimal memory may take longer or be unable to complete the tests.

```
# Within R
library(MBatch)
# Set these environment variable to override file locations if needed
#
Sys.setenv(MBATCH_TEST_OUTPUT="~/Downloads/output")
Sys.setenv(MBATCH_TEST_INPUT="~/Downloads/MATRIX_DATA")
Sys.setenv(MBATCH_TEST_COMPARE="~/Downloads/COMPARE")

baseDir <- file.path(system.file(package = "MBatch"), "tests")
message(baseDir)
testFiles <- list.files(path=baseDir)
print(testFiles)
results <- c()
for(myFile in testFiles)
{
	message("************************************************************")
	message("************************************************************")
	message("**** ", file.path(baseDir, myFile))
	message("************************************************************")
	message("************************************************************")
	test <- source(file.path(baseDir, myFile))
	if (isTRUE(test$value))
	{
		results <- c(results, paste("Test succeeded for ", myFile, sep=""))
	}
	else
	{
		results <- c(results, paste("Test failed for ", myFile, sep=""))
	}
}
print(results)
```

The last part of the output will print whether or not the tests finished properly.

```
2018 05 22 09:59:23.947 DEBUG BSTW5MHDXM2 mbatchStandardLegend before java
2018 05 22 09:59:23.978 DEBUG BSTW5MHDXM2 mbatchStandardLegend after java
No comparable output--no error means 'OK' for now.
> print(results)
 [1] "Test succeeded for AN_Adjusted.R"
 [2] "Test succeeded for AN_Unadjusted.R"
 [3] "Test succeeded for Boxplot_AllSamplesData_Structures.R"
 [4] "Test succeeded for Boxplot_AllSamplesRLE_Structures.R"
 [5] "Test succeeded for Boxplot_Group_Structures.R"
 [6] "Test succeeded for CDP_Files.R"
 [7] "Test succeeded for CDP_Plot.R"
 [8] "Test succeeded for CDP_Structures.R"
 [9] "Test succeeded for EB_withNonParametricPriors.R"
[10] "Test succeeded for EB_withParametricPriors.R"
[11] "Test succeeded for EBNPlus_CombineBatches.R"
[12] "Test succeeded for EBNPlus_Correction_Files.R"
[13] "Test succeeded for EBNPlus_Correction_Structures.R"
[14] "Test succeeded for HierarchicalClustering_Structures.R"
[15] "Test succeeded for MP_ByBatch.R"
[16] "Test succeeded for MP_Overall.R"
[17] "Test succeeded for PCA_DualBatch_Structures.R"
[18] "Test succeeded for PCA_Regular_Structures.R"
[19] "Test succeeded for RBN_Pseudoreplicates.R"
[20] "Test succeeded for RBN_Replicates.R"
[21] "Test succeeded for SupervisedClustering_Batches_Structures.R"
[22] "Test succeeded for SupervisedClustering_Pairs_Structures.R"
>
```