

The timeClust package for S-PLUS

Instructions for Installation and Use

Maintainer: Dawn Woodard, Cornell Univ.

Author: Dawn Woodard, Cornell Univ.

Overview

The timeClust package fits the time series clustering model described in Woodard and Goldszmidt (2010). The authors welcome any contributions or suggestions.

The package has been tested only on RedHat Linux, although the authors are not aware of any reason why it should not work on other platforms, including Windows. We have been unsuccessful with our attempts to port the package to R, and believe that this is due to the C++ dependencies used in the package.

timeClust includes much of the source code of the Template Numerical Toolkit, developed by the National Institute of Standards and Technology.

System Requirements

Requires S-PLUS 8.0.1+.

Instructions for Installation and Use (Linux)

1. Download the source code for the package (timeClust_0.0.tar.bz2) and unzip in a local directory using the command “tar -xvf ./timeClust_0.0.tar.bz2”
2. Edit the file installPackage in the base directory, by replacing “~/rlibs” (in both lines) with the directory where you would like to install the package.
3. Change directory using the “cd” command so that you are in the directory where you unzipped the files. Run the executable file installPackage by calling “./installPackage”. The first time you run it you will get a warning “cannot remove <lib_directory_name>: no such file or directory;” ignore this warning. After the warning it should say “Installing *source* package ‘timeClust’”, and give a list of “make”, “g++”, and “LIBRARY” statements. Then it should output some notes on the progress of the package build.
4. Open S-PLUS and call library(timeClust, lib.loc=<your_lib_dir>), where <your_lib_dir> is the directory where the timeClust package is installed.
5. Instructions for using the package and code for fitting the model to some simulated data are given in the help file for the dpMcmc function. Call help(dpMcmc) after loading the package. This help file can also be opened manually; it is located in the “man” subdirectory of the package source code. There are also help files for other functions in the package, as listed in the dpMcmc help file.

Instructions for Testing the Package (Optional)

The timeClust package includes a number of test scripts that validate the output of the model-fitting function. They can be run by changing into the same directory where the source files were unzipped, and calling “./checkPackage”. You will need to have write permissions to the current directory. Before calling checkPackage, you will need to edit the file timeClust/tests/noDataClustSize.t and noDataNclust.t and noDataZdistn.t and clustSize.t and clustSize2.t; change the directory for coda.loc to the directory where you have installed the coda package.

A warning about line endings is generated but should be ignored. After each test file is run you will get a 2-line printout giving the name of the test file and a short description of the test. If the test file passes there will be no further printout; if it fails then the code of the test file will print out. All of the tests should pass. At the end you will get a line saying “ERROR”, with no description of the error; this is a bug with S-PLUS and does not indicate that any of the tests failed. Again, as long as you only get a two-line printout for each test then the test has passed.

An additional method for checking correctness of the Markov chain output is implemented in the priorPostClust function in the package; call help(priorPostClust) for more information.

Data

The data that is included in the package has been simulated using the simulateTS function in the package; call help(simulateTS) for more information.