Instructions.

In this folder ("LabelSwitching") we have included all the code used in Subsection: *Relabeling algorithms, Gibbs Sampler.* This will enable readers to replicate all the analysis, simulations and comparisons described in a that part of the paper and to apply our proposed strategy (and alternative algorithms) to their own data.

The folder "LabelSwitching" contains the following files and folders:

1. instructions.pdf: this file!

2. gibbs: this folder contains the **C** code to implement a standard Gibbs sampler and related subroutines.

3. relabel: this folder contains the **C** code to implement our relabeling algorithm and the alternative proposals as described in the paper.

4. data_sets: this folder contains the data sets analyzed in our manuscript and other well known data sets from the mixture modeling literature.

5. results: once the readers replicate the analysis, all the results will be stored automatically in this folder.

6. Functions.r: in this **R** file are all the functions of **R** needed to replicate our experiments. Each function has been described and commented.

7. Example.r: in this **R** file we have outlined a complete analysis using our proposal and alternative algorithms.

As described in the paper, we coded our approach in **C** and used the `.C` interface to R to have a friendly data input interface, see Pend and Leew (2002) or Darren Wilkinson's (`http://darrenjw.`

`wordpress.com/2010/12/30/calling-c-code-from-r/`) research blog for a good introduction. We have done this under Linux (openSUSE and Kubuntu), so you should be able to run our code under any Unix/Linux environment.

Hence, assuming a Unix/Linux environment, our code can be compiled into a shared library following the next simple steps:

1. Open a Unix/Linux shell.

2. Go to the directory where the **C** code for the standard Gibbs sampler has been stored e.g.

    cd /home/username/Documents/LabelSwitching/gibbs

3. The **C** code in this folder can be compiled with a command like[1]:

    R CMD SHLIB comp_pred.c posterior_mean.c compar_posterior_mean.c utilities.c gibbs_normal.c main_normal.c -o univ_gibbs.so

    Now, in the file gibbs, you should have the shared library "univ_gibbs.so". If you experience any problem, please check first that you have a complete installation of the **C/C++** development pattern and second that you have installed the development package of R and not just the base version.

4. Then go to the directory where the **C** code for the relabeling strategies have been stored e.g.

    cd /home/username/Documents/LabelSwitching/relabel

5. Now the command is [2]:

    R CMD SHLIB classification.c perm_min_rel_error.c KL_initial_label.c SampleNoRep.c KL_relabel.c ECR_relabel.c group_mean_var.c data_relabel.c lsap.c assignment.c -o relabel.so

    This creates the shared library "relabel.so"

---

[1]Copy and paste instructions from a PDF to a shell is not a good idea: we included the text file (in the file gibbs) "compile_gibbs.txt" with the same command.

[2]See the file "compile_relabel.txt".

6. Using an R editor[3], open the file "Example.r". The idea is to simply modify the bits of this file that are uncommented and send instructions from the editor to **R**. For example, in the next paragraph, all the lines with # are comments, and this comments should work only as a guidance. But the line with no #, must be modified accordingly.

```
#   Directory where the folder LabelSwitching is              #
# e.g. if the path is (this was our case)                     #
#                                                             #
#   /home/carloserwin/Documents/LabelSwitching/               #
#                                                             #
# then                                                        #
#                                                             #
# dir_work <- '/home/carloserwin/Documents/LabelSwitching/'   #
#                                                             #
dir_work <- '/home/carloserwin/Desktop/LabelSwitching/'
```

7. The file "Functions.r" contains all the functions used and it has headers and numbers for easy identification with the file "Example.r". The input and output parameters of each function have been described concisely. Guidance is given to include additional data sets.

8. Because of their length, some instructions have been partitioned:

```
tt0 <- system.time(normal_gibbs(burn, m2, in_out$k, in_out$y, delta,
                          mu0, kappa, alpha, gg, hh, in_out$dir_res))[3]
```

here the two lines must be send to **R**.

9. Warnings: most of the functions depend on a previous function, i.e. if you want to generate the scaled density predictive but you have not yet generated a relabeling: **R** will crash! Run all the functions one by one, each function has an number, this does indicate precedence. All the results will be stored automatically in text files, once that you have run an analysis for

---

[3]RStudio, Kate (KDE) or the one that you like best.

a particular data set you will not need to run it again. In this case you can simply run core functions[4] and then summary functions or graphics. Our code makes direct use of directories: you must be sure where the folder "LabelSwitching" is. We have not included security checks in our functions and we offer ABSOLUTELY NO WARRANTY.

10. Please do let me know if you find any bugs in our code: Carlos E. Rodríguez, cerh2@kent.ac.uk.

REFERENCES

Peng, R. D., and Leeuw, J. , An Introduction to the `.C` Interface to R, UCLA: Academic Technology Services, Statistical Consulting Group, `http://www.ats.ucla.edu/stat/r/library/interface.pdf` (2002)

---

[4]Load the data set and set the directory where the MCMC output and relabelings are stored.