regSpec

norbert <norbert@stat.wisc.edu> Fri, Oct 17, 2014 at 1:01 PM
To: Karl Rohe <karlrohe@stat.wisc.edu>
Cc: tai qin <danielqintai@hotmail.com>, Juhee Cho <chojuhee0722@gmail.com>, Mohammad Khabbazian <khabbazian@wisc.edu>, Song Wang <songwang@stat.wisc.edu>, Thu Le <thuhale@cs.wisc.edu>, Donggyu Kim <kimd@stat.wisc.edu>, yilin zhang <yzhang@stat.wisc.edu>, Zoe Russek <zrussek@wisc.edu>, Nick Zaborek <zaborek@wisc.edu>, Xiaoyi Yang <xyang222@wisc.edu>

I think that for doing regularized spectral clustering it's better to use the rARPACK package. It has an eigs() function that is equivalent to the function in Matlab (they both use ARPACK). Last time I checked the function does not accept a sparse symmetric matrix (dsCMatrix), so you have to converted to a nonsymmetric matrix (dgCMatrix) before passing it into the function. Based on some tests I ran a while back, it is much faster than irlba.

In an effort to be more organized, I have started putting some of the functions I commonly use in a simple R package that I have up on my github. It includes a function that does regularized spectral clustering. If you want to have a look it can be found here: https://github.com/norbertbin/SpecClustPack

-Norbert

On 2014-10-17 09:51, Karl Rohe wrote:

> It is probably a good idea for us all to be using the same regularized spectral clustering algorithm. So, I have put a file regSpec.R in our shared dropbox folder. If you use matlab and you have a similar file, it would be great if you wouldn't mind sharing your file as well in the shared folder.
>
> If you have any comments on the code or if you get an error message, please share. I will plan to post this on my website in the next few months.
>
> ```
> # This code implements regularized spectral clustering for
> # symmetric, directed, and bipartite (i.e. rectangular A).
> # It has the following implemented:
> # uses irlba package
> # checks if symmetric (currently, uses svd either way)
> # uses regularization
> # project onto sphere
> # use kmeans ++
> # plots top eigenvalues and eigenvectors
> # returns projected singular vectors and kmeans++ clusters.
>
> # For symmetric calculations, it would be faster to find eigenvectors.
> #  With current R libraries, it looks like one would need to use RcppEigen
> #  to access the c++ library Eigen. Then, find the relevant function
> in that library.
>
> #  irlba uses default parameters
>
> # Version 0.1.  Oct 17, 2014.  karlrohe@stat.wisc.edu
> ```

> If you have any comments, feel free to share.