Read Me

The replication file consists the codes and helper files to replicate the Monte Carlo simulations for success of Maximal Clique and HF method. The codes can be decomposed to three steps; data generating, analyzing and evaluating. Outputs of the first two steps are files saved in disk and the last step are tables consist of evaluations.

The codes are designed in R project version 3.11 and Gauss 10 for Windows and developed in Windows 7 64bit. To run the code trouble-free, the following requirements must be satisfied.

Requirements

1. R project, Gauss 6.0 for Windows (or newer)

2. Following packages for R: "igraph", "urca", "stringr", "rugarch", "parallel"

3. Gauss should be installed in directory C:/, not any subdirectory of C:/ (since the code calls Gauss from C:/gauss6.0/tgauss)

Usage

Run all the commands in main.R. The file contains commands used for data generating, analyzing and evaluating for both method and each data type. The results are saved in Results folder as csv file. NOTE: The tables in these files are re-formed for the paper but both shares same outline.

Codes

1. dataGen.R: The file contains two codes, dataGen, dataGenplus that generates single and multiple club data respectively.

   dataGen: Script for generating data involving single club

   Inputs:

   - Tm: Time step (length) of the data,
   - n: Number of the countries included in data,
   - clubSize: Number of the club members,
   - frho: AR coefficient of the nonstationary factor's error terms,
   - trial: quantity of the data that will be produced,
   - noCons: TRUE if logGDP series is not wanted to contain intercept.

   Outputs: (Not returned but saved to hard disk.)

   - z: Matrix of n countries where clubsize of them are convergent,
   - ind: indice vector states which countries are converging,
   - gammas: vector of country factor loadings.

   dataGenplus: Script for generating data involving multiple clubs whose club sizes are selected among random Poisson numbers.

   Inputs:

   - Tm: Time step (length) of the data,
   - n: Number of the countries included in data,
   - k: Number of clubs,
   - frho: AR coefficient of the nonstationary factor's error terms,

- trial: Quantity of the data that will be produced,
- noCons: TRUE if logGDP series is not wanted to contain intercept.

Outputs: (Not returned but saved to hard disk.)

- z: Matrix of n countries form k convergence clubs
- ind: Indice vector states which countries are converging,
- gammas: Vector of country factor loadings.

2. dataGenplus_old.R: Script for generating data involving multiple clubs whose club sizes are selected equal (see the paper).

   Inputs:

   - Tm: Time step (length) of the data,
   - n: Number of the countries included in data,
   - k: Number of clubs,
   - frho: AR coefficient of the nonstationary factor's error terms,
   - trial: Quantity of the data that will be produced,
   - noCons: TRUE if logGDP series is not wanted to contain intercept.

   Outputs: (Not returned but saved to hard disk.)

   - z: Matrix of n countries form k convergence clubs
   - ind: Indice vector states which countries are converging,
   - gammas: Vector of country factor loadings.

3. mcAGK.R: The file contains the script, mcAGK, that detects clubs of data containing *single club* via maximal clique algorithm employing ADF, ADF-GLS and KPSS at 5% significance level.

   Inputs:

   - Tm: Time step (length) of the data,
   - n: Number of the countries included in data,
   - clsize: Number of club members,
   - frho: AR coefficient of the nonstationary factor's error terms,
   - noCons: TRUE if logGDP series doesn't contain intercept.

   Outputs: (Not returned but saved to hard disk.)

   - Report: Matrix containing values to evaluate success,
   - Gmml: Matrix containing gammas with detected convergent and non-convergent country lists.

4. mcHF.R: The file contains the script, mcHF, that detects both perfect and relative convergence clubs of data containing *single club* via HF algorithm at 5% significance level.

   Inputs:

   - Tm: Time step (length) of the data,
   - n: Number of the countries included in data,
   - clsize: Number of club members,
   - frho: AR coefficient of the nonstationary factor's error terms,
   - noCons: TRUE if logGDP series doesn't contain intercept.

Outputs: (Not returned but saved to hard disk.)

- Report: Matrix containing values to evaluate success,
- Gmml: Matrix containing gammas with detected convergent and non-convergent country lists.

5. mcAplus.R: The file contains the script, mcAplus, that detects clubs of data containing *multiple clubs* via maximal clique algorithm employing ADF at 5% significance level.

Inputs:

- Tm: Time step (length) of the data,
- n: Number of the countries included in data,
- k: Number of clubs,
- frho: AR coefficient of the nonstationary factor's error terms,
- noCons: TRUE if logGDP series doesn't contain intercept.
- nopois: TRUE if data is generated by dataGenplus_old; dataGenplus otherwise.

Outputs: (Not returned but saved to hard disk.)

- Report: Matrix containing values to evaluate success,
- Gmml: Matrix containing gammas with detected convergent and non-convergent country lists.

- Report: Matrix containing values to evaluate success,
- Gmml: Matrix containing gammas with detected convergent and non-convergent country lists.

6. mcHFplus.R: The file contains the script, mcHFplus, that detects clubs of data containing *multiple clubs* via HF algorithm at 5% significance level.

Inputs:

- Tm: Time step (length) of the data,
- n: Number of the countries included in data,
- k: Number of clubs,
- frho: AR coefficient of the nonstationary factor's error terms,
- noCons: TRUE if logGDP series doesn't contain intercept.
- nopois: TRUE if data is generated by dataGenplus_old; dataGenplus otherwise.

Outputs: (Not returned but saved to hard disk.)

- Report: Matrix containing values to evaluate success,
- Gmml: Matrix containing gammas with detected convergent and non-convergent country lists.

7. anlysSingle.R: The file contains the code overallRep (with its helper functions) that evaluates single club results for both cases of intercept and for both methods.

Inputs: (no input).
Outputs: Returns following evaluation tables (saved in Results Folder):

- KSrepNoConst: Table consisting Hit Rates, False Alarm Rates and Kupiers Scores for all data type and both method,
- PTSnoConst: Table consisting Hit Rates, False Alarm Rates and Kupiers Scores for all data type and both method,
- KSrepWithConst: Table consisting Hit Rates, False Alarm Rates and Kupiers Scores for all data type and both method,

- PTSwithConst: Table consisting Hit Rates, False Alarm Rates and Kupiers Scores for all data type and both method.

8. anlysMulti.R: The file contains the code overallRepMulti (with its helper functions) that evaluates multiple club results for both cases of intercept, both cases of club size strategy and for both methods.

   Inputs: (no input).
   Outputs: Table consisting Perfect Success results for all data type and both method. (Saved in Results Folder)

9. applyMCL.R: The file contains script for applying Maximal Clique Algorithm applying ADF to 7 different data (Penn, Maddison, 1930, Europe, Europe+G7, Europe+S&P, G7+S&P). Inputs: (no input).
   Outputs: Returns following evaluation tables (saved in Results Folder):

   - countsAllMCL.csv: Number of clubs per sizes in csv format
   - countsAllMCL.rda: Number of clubs per sizes in rda format (R file).

10. applyHF.R: The file contains script for applying HF algorithm to 7 different data (Penn, Maddison, 1930, Europe, Europe+G7, Europe+S&P, G7+S&P). Outputs: Returns following evaluation tables (saved in Results Folder):

    - countsAllHF.csv: Number of clubs per sizes in csv format.