Changes in RSiena for implementing the GMoM

Viviana Amati

February 2021

This file reports all the changes that have been implemented in `RSiena` for the GMoM estimation. Changes for the Robbins-Monro algorithm concern mainly the R code, while changes for the definition of the statistics concern the C++ code.

Folder Data

Basic GMoM effects have been included in the file `alleffects.csv`. They are fixed by default to 0 because the GMoM statistics are only used to evaluate the moment conditions and are not part of the probability model.

Folder src

- Added the files for the computation of the statistics for the network evolution only in the folder model/effects
  (Files: `AgreementTransitivityGMMEffect.h, RealTransitivityGMMEffect.h`, `ReciprocityGMMEffect.h`)
- Files Model.cpp, Model.h, siena07internals.cpp, Stattisticscalculator.cpp: declaration of the effect type "gmm" and computation of the GMoM statistics
- `model/AllEffects.h`: inclusion of the gmom effects for the network evolution only
- `model/EffectFactory.h, model/EffectFactory.cpp`: implementation of the effect of type gmm, change short name of the statistics from simX_sim to simX_gmm
- `model/AverageAlterEffect.h, model/AverageAlterEffect.cpp`: implementation of the GMoM statistic averageAlter
- `model/MaxAlterEffect.h, model/MaxAlterEffect.cpp`: implementation of the GMoM statistics maxAlter and minAlter
- `model/SimilarityEffect.h, model/SimilarityEffect.cpp`: implementation of the GMoM statistic simX

Folder R

- `effects.R`: replacement text for the GMoM statistics
- `effectsMethods.R`: in function `print.sienaEffects`, when one or more GMoM statistics are included, and the effects object is printed in the console, both model effects and statistics are shown.

```
 Effects and statistics for estimation by the Generalized Method of Moments

 Effects
   name  effectName                      include fix   test  initialValue parm type
1  mynet constant mynet rate (period 1) TRUE    FALSE FALSE    4.69604   0    rate
2  mynet constant mynet rate (period 2) TRUE    FALSE FALSE    4.32885   0    rate
3  mynet outdegree (density)             TRUE    FALSE FALSE   -1.46770   0    eval
4  mynet reciprocity                     TRUE    FALSE FALSE    0.00000   0    eval
5  mynet transitive triplets             TRUE    FALSE FALSE    0.00000   0    eval
6  mynet 3-cycles                        TRUE    FALSE FALSE    0.00000   0    eval
7  mynet mybeh ego                       TRUE    FALSE FALSE    0.00000   0    eval
8  mynet mybeh similarity                TRUE    FALSE FALSE    0.00000   0    eval
9  mybeh rate mybeh (period 1)           TRUE    FALSE FALSE    0.70571   0    rate
10 mybeh rate mybeh (period 2)           TRUE    FALSE FALSE    0.84939   0    rate
11 mybeh mybeh linear shape              TRUE    FALSE FALSE    0.32237   0    eval
12 mybeh mybeh quadratic shape           TRUE    FALSE FALSE    0.00000   0    eval
13 mybeh mybeh total similarity          TRUE    FALSE FALSE    0.00000   0    eval

 Regular and GMoM statistics
   name  effectName                      Statistic
1  mynet constant mynet rate (period 1) Regular
2  mynet constant mynet rate (period 2) Regular
3  mynet outdegree (density)             Regular
4  mynet reciprocity                     Regular
5  mynet transitive triplets             Regular
6  mynet 3-cycles                        Regular
7  mynet mybeh ego                       Regular
8  mynet mybeh similarity                Regular
9  mybeh rate mybeh (period 1)           Regular
10 mybeh rate mybeh (period 2)           Regular
11 mybeh mybeh linear shape              Regular
12 mybeh mybeh quadratic shape           Regular
13 mybeh mybeh total similarity          Regular
14 mynet mybeh ego                       GMoM
15 mynet mybeh similarity                GMoM
```

– `sienaEffects.R`:
Inclusion of the function `includeGMoMStatistics`. The function allows specifying the GMoM statistics. GMoM statistics are implemented as effects of type "gmm", with fixed value 0. In this way, they are not part of the probability model but are used to evaluate the moment condition. The documentation of the function is in the file `includeGMoMStatistics.Rd`.

– `effectsDocumentation.R`:
Changes so that the GMoM statistics do not appear in the list of possible effects for the evaluation/creation/endowment functions.

– `initializeFRAN.R`:
Created objects for GMoM statistics. Prevented the diagonalization of the scaling matrix $D^{-1}B$ in the Robbins-Monro step. Defined the requested effect when preVans is used with GMoM. Adjusted code for conditional estimation.

– `phase1.R`:
The GMoM estimate for θ is the value $\hat{\theta}$ such that

$$B\ E_\theta\left[\, s^*(X,Z) - s^*(x,z)\,\right] = 0 \quad ,$$

The changes in the code relate to the computation of

- matrix of first-order derivatives of the statistics

$$\Gamma \;=\; \frac{\partial}{\partial\theta} E_\theta[s(X,Z) - s(x,z)]$$

- matrix of GMoM weights

$$W = \mathrm{Cov}[\mathrm{s(X,Z)}]$$

- matrix

$$B = \frac{\partial}{\partial\theta}\; E_\theta\,[\,s^*(X,Z) - s^*(x,z)\,]'\,W = \Gamma\; W$$

- matrix $D^{-1}B$ which is equivalent to the matrix D^{-1} in the Newton-Raphson/Robbins-Monro step for the regular MoM, i.e.

  Regular MoM:
  $\theta_{i+1} = \theta_i - \underbrace{D^{-1}}_{\text{dinvv}} E[S-s]$
  $\theta_{i+1} = \theta_i - \alpha\underbrace{D^{-1}}_{\text{dinvv}}(S-s)$

  GMoM:
  $\theta_{i+1} = \theta_i - \underbrace{D^{-1}B}_{\text{dinvv}} E[S-s]$
  $\theta_{i+1} = \theta_i - \alpha\underbrace{D^{-1}B}_{\text{dinvv}}(S-s)$

  This equivalence reduces the number of changes in phase 2, though now dinvv is not the inverse of the first order derivative matrix of the statistics anymore.

- `phase2.R`:
  The GMoM statistics are fixed at 0 by default. Thus, all the conditions involving `z$fixed` are changed into `z$fixed & !z$gmmEffects` so that they apply to the GMoM statistics as well. The term `z$gmmEffects` is a logical variable taking value TRUE if the effect is a GMoM statistic, and FALSE otherwise.

  Similar changes are required to compute the change step and set it to 0 for the GMoM statistics.

- `phase3.R`:

  - Added message: "Generalized Method of Moments estimation" for printing the results
  - Computation of the matrices: Γ, W and B as described in `phase1.r`
  - Computation of the t-ratios for convergence and tmax

$$t - \mathrm{ratio} = \frac{\hat{\mathrm{B}}(\mathrm{s(X,Z)} - \mathrm{s(x,z)})}{\sqrt{\left(\hat{\mathrm{B}}\;\hat{\Sigma}\;\hat{\mathrm{B}}'\right)}}\;.$$

- Computation the covariance matrix of the GMoM estimator as

$$\Sigma_{\widehat{\theta}_{GMoM}} = (\hat{B}\ \hat{\Gamma})^{-1}(\hat{B}\ \hat{\Sigma}\ \hat{B}')((\hat{B}\ \hat{\Gamma})^{-1})'$$

- `robmon.R`:
  - Creation of a binary variable `z$gmm` indicating whether the GMoM is used.
    - If gmm=TRUE and no gmm effects are included, the algorithm stops and asks the user to select the regular MoM estimation
    - If gmm=FALSE and gmm effects are included, the algorithm stops and asks the user to select the GMoM estimation
  - Specification of the number of steps in phase 1 for the GMoM ($n1 = 100 + 7 * p$) with p the number of parameters
- `print07Report.R`:
  Format the GMoM statistics and standard errors for printing in the Rconsole.
- `printDataReport.R`:
  Formatting for the GMoM statistics
- `sienaModelCreate.R`:
  - Include the argument `z$gmm`=TRUE for the GMoM estimation
  - Dolby is not implemented for the GMoM estimation
- `sienaprint.R`:
  Changes for printing the GMoM results in the console (theta, standard error)
- `sienatable.R`:
  Changes for printing the GMoM results in html or tex

Folder man

- `sienaAlgorithmCreate.Rd`:
  Inclusion of the argument gmm for the generalized method of moments and corresponding documentation.
- `includeGMoMStatistics.Rd`:
  Documentation for the new function `includeGMoMStatistics`

Folder tests

`parallel.R`:
Inclusion of a test for the GMoM implementation (test15).

List of GMoM effects implemented

The GMoM estimator uses more statistics than parameters and is computed by minimizing the distance between observed statistics and their expected values. Several GMoM statistics have been implemented for the evolution of networks and co-evolution of networks and behaviors. We denote by X the network and by Z the behavior.

Evolution	Formula of the statistic
Network evolution	
New reciprocity	$s^*(X) = \sum_m \sum_{ij} (1 - X_{ij}(t_{m-1})) \cdot (1 - X_{ji}(t_{m-1})) \cdot X_{ij}(t_m) \cdot X_{ji}(t_m)$
Persistent reciprocity	$s^*(X) = \sum_m \sum_{ij} X_{ij}(t_{m-1}) \cdot X_{ji}(t_{m-1}) \cdot X_{ij}(t_m) \cdot X_{ji}(t_m)$
Real reciprocity	$s^*(X) = \sum_m \sum_{ij} (1 - X_{ij}(t_{m-1})) \cdot X_{ji}(t_{m-1}) \cdot X_{ij}(t_m) \cdot X_{ji}(t_m)$
Real Transitivity	$s^*(X) = \sum_m \sum_{ijh} (1 - X_{ij}(t_{m-1})) \cdot X_{ih}(t_{m-1}) \cdot X_{hj}(t_{m-1}) \cdot X_{ij}(t_m) \cdot X_{ih}(t_m) \cdot X_{hj}(t_m)$
Agree transitivity	$s^*(X) = \sum_m \sum_{ijh} (1 - X_{ij}(t_{m-1})) \cdot X_{ih}(t_{m-1}) \cdot X_{jh}(t_{m-1}) \cdot X_{ij}(t_m) \cdot X_{ih}(t_m) \cdot X_{jh}(t_m)$
Network-behavior evolution	
Covariate ego	$s^*(X, Z) = \sum_m \sum_i Z_i(t_m) \sum_j X_{ij}(t_m)$
Total similarity	$s^*(X, Z) = \sum_m \sum_{ij} X_{ij}(t_m)(\mathrm{sim}_{ij}^Z(t_m) - \widehat{\mathrm{sim}^Z})$
Average similarity	$s^*(X, Z) = \sum_m \sum_{ij} X_{ij}(t_m)(\mathrm{sim}_{ij}^Z(t_m) - \widehat{\mathrm{sim}^Z}) / \big(\sum_j X_{ij}(t_m)\big)$
Total alter	$s_i^{\mathrm{beh}}(X, Z) = \sum_m \sum_i Z_i(t_m) \big(\sum_j X_{ij}(t_m)\, Z_j(t_m)\big)$
Average alter	$s_i^{\mathrm{beh}}(X, Z) = \sum_m \sum_i Z_i(t_m) \big(\sum_j X_{ij}(t_m)\, Z_j(t_m)\big) / \big(\sum_j X_{ij}(t_m)\big)$
Minimum alter	$s_i^{\mathrm{beh}}(X, Z) = \sum_m \sum_i Z_i(t_m) \big(\min_j\ X_{ij}(t_m)\, Z_j(t_m)\big)$
Maximum alter	$s_i^{\mathrm{beh}}(X, Z) = \sum_m \sum_i Z_i(t_m) \big(\max_j\ X_{ij}(t_m)\, Z_j(t_m)\big)$

Table 1: Implemented GMoM statistics and corresponding formulas

The statistics for the network evolution complement the information provided by the regular statistics by considering two consecutive observations of the network, i.e., the statistics depend on both the network at time t_m and that at time t_{m-1}, $s^*(X(t_m), X(t_{m-1}))$. The GMoM statistics for the co-evolution of networks and behaviors supplement the information of the cross-lagged statistics depending jointly on the network and behavior by considering the network and behavior at the same time point, i.e., $s^*(X(t_m), Z(t_m))$.

Essentially, the implemented effects are those in the two published papers and a few more. The complete list and corresponding definitions are describe in Table 1.

Other effects can be implemented following the procedure described in the next section.

How to implement a gmom effect

The following lines explain how to implement a GMoM effect. The code for the `egoX_gmm` statistic is used as an illustration.

1. Start from the CPP and H files of the regular statistic corresponding to the GMoM statistic that one wants to implement. Those files are in the folder `src/model/effects`.

2. In the file .CPP duplicate the constructor and include the `simulatedState` flag. When the `simulatedState` flag is set to TRUE the value of the behaviour or the network is taken at the end of the period.

```
CovariateEgoEffect::CovariateEgoEffect(const EffectInfo * pEffectInfo,
const bool leftThresholded, const bool rightThresholded,
const bool simulatedState) :
CovariateDependentNetworkEffect(pEffectInfo, simulatedState) {
this->lleftThresholded = leftThresholded;
this->lrightThresholded = rightThresholded;
this->lthreshold = pEffectInfo->internalEffectParameter();
}
```

3. Declare the `simulatedState` flag in the file .H

```
CovariateEgoEffect(const EffectInfo * pEffectInfo,
const bool leftThresholded, const bool rightThresholded,
const bool simulatedState);
```

4. Include the new effect in the file `EffectFactory.cpp`.

```
else if (effectName == "egoX_gmm")
{
 pEffect = new CovariateEgoEffect(pEffectInfo, false, false, true);
}
```

5. Include the new effect in the file `allEffects.csv`. This can be done by copying the line of the regular statistic, changing the short name by adding the suffix `_gmm`, and setting the value of the columns type and fix to “gmm” and “TRUE”, respectively.