Generalized Empirical Likelihood with R

Pierre Chaussé

November 11, 2026

Abstract

This an extention of the gmmS4 vignette, to explain how to use the package for generalized empirical likelihood estimation.

Contents

1 A very brief review of the GEL method **3**
1.1 An S4 class object for moment based models 4
1.2 The rhoXXX functions 5
1.3 The Lagrange multiplier solver 5

2 The *solveGel* Method **6**

3 The *gelFit* Method **8**
3.1 Methods for "gelfit" classes 9
3.2 Robust to misspecification standard errors 14

4 The *evalGel* Method **15**

5 One function to fit them all: gel4 **16**

6 GEL confidence intervals for the mean **18**

A Some extra codes **20**

1 A very brief review of the GEL method

We present how to use the package to estimate models by the Generalized Empirical Likelihood method (GEL) (see Newey and Smith (2004) for the iid case and Anatolyev (2005) for weakly dependent processes). We assume that the reader has read the gmmS4 vignette in which many classes and methods needed are defined. We first describe the method without going into too much details. The author can refer to the above papers for a detailed description, or Chaussé (2010) who explains GEL estimation using the gmm (Chaussé (2019))package.

The estimation is based on the following moment conditions

$$\mathrm{E}[g_i(\theta)] = 0,$$

For the iid case, the estimator is defined as the solution to either

$$\hat{\theta} = \arg\min_{\theta, p_i} \sum_{i=1}^{n} h_n(p_i)$$

subject to,

$$\sum_{i=1}^{n} p_i g_i(\theta) = 0$$

and

$$\sum_{i=1}^{n} p_i = 1,$$

where $h_n(p_i)$ belongs to the following Cressie-Read family of discrepancies:

$$h_n(p_i) = \frac{[\gamma(\gamma+1)]^{-1}[(np_i)^{\gamma+1} - 1]}{n},$$

or

$$\hat{\theta} = \arg\min_{\theta} \left[\max_{\lambda} \frac{1}{n} \sum_{i=1}^{n} \rho\left(\lambda' g_i(\theta)\right) \right] \tag{1}$$

The first is the primal and the second is the dual problem, the latter being preferred in general to define GEL estimators. The vector λ is the Lagrange multiplier associated with the first constraint in the primal problem. Its estimator plays an important role in testing the validity of the moment conditions. $\rho(v)$ is a strictly concave function normalized so that $\rho'(0) = \rho''(0) = -1$. It can be shown that $\rho(v) = \ln(1-v)$ corresponds to Empirical Likelihood (EL) of Owen (2001) , $\rho(v) = -\exp(v)$ to the Exponential Tilting (ET) of Kitamura and Stutzer (1997), and $\rho(v) = (-v - v^2/2)$ to the Continuous Updated GMM estimator (CUE) of Hansen et al. (1996). In the context of GEL, the CUE is also known at the Euclidean Empirical Likelihood (EEL), because it corresponds to $h_n(p_i)$ being the Euclidean distance.

Once the solution is obtained for θ and λ, the implied probabilities can be computed as follows

$$\hat{p}_i = \frac{\rho'(\hat{\lambda}' g_i(\hat{\theta}))}{\sum_{j=1}^{n} \rho'(\hat{\lambda}' g_j(\hat{\theta}))} \tag{2}$$

If we relax the iid assumption, the problem is identical, but the moment function must be smoothed using a kernel method. Smith (2011) proposes to replace $g_i(\theta)$ by:

$$g_i^w(\theta) = \sum_{s=-m}^{m} w(s) g_{i-s}(\theta)$$

where $w(s)$ are kernel based weights that sum to one (see also Kitamura and Stutzer (1997) and Smith (2011).

1.1 An S4 class object for moment based models

There is no dinstrinction between GEL and GMM models. They are all objects of class "momentModel", and they are described in the gmmS4.pdf vignette. For the purpose of presenting the GEL method, we create the same models.

- Linear models:

```
library(momentfit)
data(simData)
lin <- momentModel(y~x1+x2, ~x2+z1+z2, data=simData, vcov="iid")
```

- Formula-type models:

```
theta0=c(mu=1,sig=1)
x <- simData$x1
dat <- data.frame(x=x, x2=x^2, x3=x^3, x4=x^4)
gform <- list(x~mu,
              x2~mu^2+sig,
              x3~mu^3+3*mu*sig,
              x4~mu^4+6*mu^2*sig+3*sig^2)
form <- momentModel(gform, NULL, theta0=theta0, vcov="MDS", data=dat)
form

## Model based on moment conditions
## *********************************
## Moment type: formula
## Covariance matrix: MDS
## Number of regressors: 2
## Number of moment conditions: 4
## Number of Endogenous Variables: 0
## Sample size:  50
```

- function:

```
fct <- function(theta, x)
    cbind(x-theta[1], (x-theta[1])^2-theta[2],
          (x-theta[1])^3, (x-theta[1])^4-3*theta[2]^2)
dfct <- function(theta, x)
    {
        m1 <- mean(x-theta[1])
        m2 <- mean((x-theta[1])^2)
        m3 <- mean((x-theta[1])^3)
        matrix(c(-1, -2*m1, -3*m2, -4*m3,
                 0, -1, 0, -6*theta[2]), 4, 2)
    }
theta0=c(mu=1,sig2=1)
func <- momentModel(fct, simData$x3, theta0=theta0, grad=dfct, vcov="iid")
func

## Model based on moment conditions
## *********************************
## Moment type: function
## Covariance matrix: iid
## Number of regressors: 2
## Number of moment conditions: 4
## Number of Endogenous Variables: 0
## Sample size:  50
```

- Non-linear models:

```
theta0=c(b0=1, b1=1, b2=1)
gform <- y~exp(b0+b1*x1+b2*x2)
nlin <- momentModel(gform, ~z1+z2+z3+x2, theta0=theta0, vcov="MDS", data=simData)
nlin

## Model based on moment conditions
## *********************************
## Moment type: nonlinear
## Covariance matrix: MDS
## Number of regressors: 3
## Number of moment conditions: 5
## Number of Endogenous Variables: 2
## Sample size:  50
```

1.2 The rhoXXX functions

The package provides $\rho(v)$ for ET, EL, EEL and HD. The function names are respectively "rhoET", "rhoEL", "rhoEEL" and "rhoHD". For any other GEL, users can pass user specific $\rho(v)$ function to the rhoFct argument of the gelModel(). The following example shows how the function must be built:

```
rhoEL

## function (gmat, lambda, derive = 0, k = 1)
## {
##     lambda <- c(lambda) * k
##     gmat <- as.matrix(gmat)
##     gml <- c(gmat %*% lambda)
##     switch(derive + 1, log(1 - gml), -1/(1 - gml), -1/(1 - gml)^2)
## }
## <bytecode: 0x606d2227acf0>
## <environment: namespace:momentfit>
```

Therefore, the function must be in the form $f(X, \lambda, d, k)$, where the first argument is an $n \times q$ matrix, the second argument is a vector of q elements, the third is an integer that indicates the order of derivative to return, and the last is a scalar that depends on the kernel used to smooth the moment function. We will discuss that in more details below. The function must return the vector $\rho(kX\lambda)$, $\rho'(kX\lambda)$ or $\rho''(kX\lambda)$, when derive is 0, 1, or 2, where $\rho'(v)$ and $\rho''(v)$ are the first and second derivative of $\rho(v)$.

1.3 The Lagrange multiplier solver

The function getLambda() function solve the maximization problem

$$\max_{\lambda} \frac{1}{n} \sum_{i=1}^{n} \rho(\lambda' X_i)$$

It is used to solve problem (1). The arguments are:

```
args(getLambda)

## function (gmat, lambda0 = NULL, gelType = NULL, rhoFct = NULL,
##     tol = 1e-07, maxiter = 100, k = 1, method = "BFGS", algo = c("nlminb",
##         "optim", "Wu"), control = list(), restrictedLam = integer(),
##     ...)
## NULL
```

The first argument is the $n \times q$ matrix X. The second argument is the starting value for λ. If set to NULL, a vector of zeroes is used. The third argument is the GEL type, which is either "ET", "EL",

"EEL", "REEL" or "HD". If the rhoFct is provided, getType is ignored. The argument control is used to pass a list of arguments to optim(), nlminb() or constrOptim(). The argument "k" is the scalar described in the previous subsection. To describe all other arguments, we are presenting the options by type of GEL:

- EL: There are three possible options for EL. The default is "nlminb", in which case, the arguments tol, maxiter and method are ignored. If algo is set to "optim", the solution is obtained using constrOptim(), with the restriction $(1 - \lambda' X_i) > 0$ for all i. If algo is set to "Wu", the Wu (2005) algorithm is used. The argument tol and maxit are used to control the stopping rule.
- HD: Identical to EL, except that the "Wu" algorithm is not available for that type of GEL.
- EEL: There is an analytical solution, so all other arguments are ignored.
- REEL: This is EEL with a non-negativity constraint on all implied probabilities. In that case, the maxiter can be used to control the number of iterations.
- Others: When rhoFct is provided or the type is ET, the solution is obtained by either "nlminb" or "optim". In that case, the algorithms are controlled through the control argument.

Here are a few example using the simulated dataset (convergence code of 0 means that the algorithm converged with no detected problem):

```
X <- simData[c("x3","x4","z5")]
(res <- getLambda(X, gelType="EL"))$lambda

## [1] -0.04178415 -0.06745195 -0.02075063

res$convergence$convergence

## [1] 0
```

```
(res <- getLambda(X, gelType="ET",control=list(maxit=2000)))$lambda

## [1] -0.04171694 -0.06804520 -0.02120255

res$convergence$convergence

## [1] 0
```

The following shows that we can provide getLambda() with a rhoFct instead:

```
(res <- getLambda(X, rhoFct=rhoEEL))$lambda

## [1] -0.04119158 -0.06769691 -0.02148683

res$convergence$convergence

## [1] 0
```

Although we rarely call the function directly, it is good to understand how it works, because it is an important part of the estimation procedure. @

2 The *solveGel* Method

The main method to estimate a model by GEL methods is *solveGel*. The available signatures are:

```
showMethods("solveGel")

## Function: solveGel (package momentfit)
## object="momentModel"
```

The arguments are

- theta0: The initial value for the minimization algorithm. It is required if the model does not have a theta0 slot. If it has a theta0 slot, it is used by default unless a new theta0 is provided.
- lambda0: The initial value to pass to the lambda solver. By default, it is set to 0, which is its asymptotic value in case of correctly specified models.
- lamSlv: An optional function to solve for lambda. By default, getLambda() is used. The function must have the following form:

```
mylSolve <- function(gmat, lambda0, gelType=NULL, rhoFct=NULL, k=1, ...)
    {
        lambda <-  rep(0,ncol(gmat))
        obj <-  sum(colMeans(gmat)^2)
        list(lambda=lambda, convergence=0, obj=obj)
    }
```

Therefore, it must return a list with lambda, convergence and obj. In the above example, λ is set to a vector of zeros and the returned obj is $\bar{g}(\theta)'\bar{g}(\theta)$. The solution will therefore be the one step GMM with the weighting matrix equals to the identity.

```
solveGel(lin,theta0=c(0,0,0), lamSlv=mylSolve)

## $theta
##  (Intercept)           x1           x2
## -0.226153103  1.009952150  0.002381384
##
## $convergence
## [1] 0
##
## $lambda
## (Intercept)          x2          z1          z2
##           0           0           0           0
##
## $lconvergence
## [1] 0
##
## $restrictedLam
## integer(0)
```

To present a more realistic example, suppose we want to estimate the model using the exponentially tilted empirical likelihood (ETEL) method of Schennach (2007), we could write a function that solves the lambda using ET, and returns the empirical log-likelihood ratio:

```
mylSolve <- function(gmat, lambda0=NULL, gelType=NULL, rhoFct=NULL, k=1, ...)
    {
        gmat <- as.matrix(gmat)
        res  <-  getLambda(gmat, lambda0, gelType="ET", k=k)
        gml <- c(gmat%*%res$lambda)
        w <- exp(gml)
        w <- w/sum(w)
        n <- nrow(gmat)
        res$obj <- mean(-log(w*n))
        res
    }
etelFit <- solveGel(lin,theta0=c(1,1,0), lamSlv=mylSolve)
etelFit$theta

## (Intercept)          x1          x2
##   1.3266917   0.8504703  -0.1030545
```

That's equivalent to setting gelType to "ETEL":

```
solveGel(update(lin, gelType="ETEL"), theta0=c(1,1,0))$theta

## (Intercept)          x1          x2
##   1.3272620   0.8505003  -0.1031604
```

- coefSlv: A character string that indicates the name of the minimization solver used to obtain $\hat{\theta}$. By default, "optim" is use. The other options are "nlminb" and "constrOptim".
- lControl: A list of options for the lambda solver. It is passed to getLambda() or lamSlv() if provided. By default, the Wu (2005) method is used to solve for the Lagrange multiplier. For example, we can use nlminb() instead:

```
solveGel(lin, theta0=c(1,1,0), lControl=list(algo="nlminb"))$theta

## (Intercept)          x1          x2
##   1.3272620   0.8505003  -0.1031604
```

It is also possible to restrict some λ's to be equal to zero. This is like removing the associated moment conditions, so the degrees of freedom of the specification tests are adjusted accordingly.

```
res <- solveGel(lin, theta0=c(1,1,0), lControl=list(restrictedLam=c(2L,3L)))
res$lambda

##   (Intercept)            x2            z1            z2
##  0.0001456767  0.0000000000  0.0000000000 -0.0001106942
```

The argument restrictedLam is a vector of integers specifying which λ to set to 0. Of course, it is an argument to use with caution. If a coefficient is only present in the moment conditions associated with the λ's set to zero, their estimates are not reliable because they no longer affect the value of the objective function.

- tControl: A list of control for the coefSlv function. We could, for example, use the following options:

```
solveGel(lin, theta0=c(1,1,0),
         tControl=list(method="BFGS", control=list(maxit=2000, reltol=1e-9)))$theta

## (Intercept)          x1          x2
##   1.3271920   0.8505091  -0.1031434
```

In that particular case, the list is directly passed to optim().

The method returns a list with: theta=$\hat{\theta}$, lambda=$\hat{\lambda}$, convergence = convergence message and code for θ, and lconvergence = convergence message and code for λ.

3 The *gelFit* Method

This is the main estimation method. It returns an object of class "gelfit". The arguments are:

- object: Any object that belongs to the union class "gelModels"
- gelType: Optional type of GEL. By default, the model is estimated by empirical likelihood.
- rhoFct: Optional $\rho(v)$ function if we want to estimate the model with a function not included in the package.
- initTheta: Method to obtain the starting value for θ. By default, the one step GMM is used. The other option is to use the one included in the object.
- theta0: The stating value for θ. If provided, the argument initTheta is ignored.
- lambda0: Starting value for λ. By default, a vector of zeros is used.

- vcov: Logical argument that specifes if the covariance matrices of $\hat{\theta}$ and $\hat{\lambda}$ should be computed? It is FALSE by default.
- ... : Additional argument that is passed to the *gelSolve* method.

In general, it works fine with the default arguments:

```
fit <- gelFit(lin)
```

The following slots are available for the "gelfit" object:

```
showClass("gelfit")

## Class "gelfit" [package "momentfit"]
##
## Slots:
##
## Name:          theta   convergence        lambda  lconvergence          call
## Class:       numeric       numeric       numeric       numeric    callORNULL
##
## Name:        gelType          vcov         model restrictedLam      argsCall
## Class:          list          list   momentModel       integer          list
```

3.1 Methods for "gelfit" classes

- *print*: It prints the model, $\hat{\theta}$ and $\hat{\lambda}$. The arguments "model" and "lambda" can be set to FALSE to avoid printing them.

```
fit <- gelFit(lin)
print(fit, lambda=FALSE, model=FALSE)

##
## Estimation:  EL
## Convergence Theta:  0
## Convergence Lambda:  0
## coefficients:
## (Intercept)           x1           x2
##   1.3270715    0.8505249   -0.1031374
```

- *residuals*: For "linearGel" and "nonlinearGel", it returns the difference between the left hand side and right hand side.
- *getImpProb*: It computes the vector of implied probabilities $\hat{p}_i$'s. By default, they are defined as:

$$\hat{p}_i = \frac{-\rho'(\hat{\lambda}' g_i(\hat{\theta}))/n}{\sum_{j=1}^{n} -\rho'(\hat{\lambda}' g_j(\hat{\theta}))/n}$$

The options are

  - posProb: This option is only considered for EEL type of GEL. If set to TRUE, the method of Antoine et al. (2007) is used to compute the probabilities:

$$\tilde{p}_i = \frac{\hat{p}_i + \varepsilon/n}{1+\varepsilon},$$

where $\varepsilon = -n \min(\min_i(\hat{p}_i), 0)$, and $\hat{p}_i = -\rho'(g_i(\hat{\theta}))/n$, which is the unormalized version of the above definition. It is then normalized to sum to 1. The method is only needed when we want to compute moments based on them. It is therefore set to FALSE by default.
  - normalize: If TRUE, the default, the above normalized version is returned. If FALSE, it returns $-\rho'(g_i(\hat{\theta}))/n$.

The method returns a list with pt, the vector $\{\hat{p}_i\}$, convMom $= \sum_{i=1}^{n} \hat{p}_i g_i(\hat{\theta})$, and convProb $= |(\sum_{i=1}^{n} \hat{p}_i) - 1|$. convMom is important, because it indicates if the estimation went well. It should be a vector close to zero.

```
pt <- getImpProb(fit)
pt$convMom

##   (Intercept)            x2            z1            z2
## -4.702727e-18  1.251169e-16  2.916504e-17 -3.005951e-17
```

- *vcov*: It returns the covariance matrices of $\hat{\theta}$ and $\hat{\lambda}$ in a list. The returned covariance matrix for $\hat{\theta}$ is:

$$\hat{\Sigma} = \frac{1}{n}\left[\hat{G}'\hat{\Omega}^{-1}\hat{G}\right]^{-1}$$

when the "vcov" component of the object is not "HAC", with

$$\hat{G} = \frac{1}{n}\sum_{i=1}^{n}\frac{dg_i(\hat{\theta})}{d\theta},$$

and

$$\hat{\Omega} = \frac{1}{n}\sum_{i=1}^{n} g_i(\hat{\theta})g_i(\hat{\theta})'.$$

For HAC specification, the covariance matrix is:

$$\hat{\Sigma}_w = \frac{1}{n}\left[\hat{G}_w'\hat{\Omega}_w^{-1}\hat{G}_w\right]^{-1},$$

with

$$\hat{G}_w = \frac{1}{nk_1}\sum_{i=1}^{n}\frac{dg_i^w(\hat{\theta})}{d\theta}$$

and

$$\hat{\Omega}_w = \frac{b}{nk_2}\sum_{i=1}^{n} g_i^w(\hat{\theta})g_i^w(\hat{\theta})',$$

where b is the bandwidth and the scalars k_1 and k_2 are determined by the kernel. The values of $\{k_1, k_2\}$ are $\{2, 2\}$ when the smoothing is based on the Truncated kernel, which implies Bartlett for the HAC estimator, and $\{1, 2/3\}$ when the smoothing is based on the Bartlett kernel, which implies Parzen for the HAC estimator.
Using the above definitions, the returned covariance matrix for $\hat{\lambda}$ is

$$\widehat{Var(\hat{\lambda})} = \frac{1}{n}\left[\hat{\Omega}^{-1} - \hat{\Omega}^{-1}\hat{G}\hat{\Sigma}\hat{G}'\hat{\Omega}^{-1}\right]$$

for non-smoothed moments, and

$$\widehat{Var(\hat{\lambda})}_w = \frac{b^2}{n}\left[\hat{\Omega}_w^{-1} - \hat{\Omega}_w^{-1}\hat{G}_w\hat{\Sigma}_w\hat{G}_w'\hat{\Omega}_w^{-1}\right]$$

for smoothed moments. The method has two additional arguments. The first is "withImpProb", which is FALSE by default. if set to TRUE, all moment estimations in the form $\sum_{i=1}^{n}[]_i/n$ are replaced by $\sum_{i=1}^{n}\hat{p}_i[]_i$. The second argument is tol, which is a tolerance level for too small

diagonal elements of the covariance matrix of $\hat{\lambda}$. When some moment conditions are not binding, the estimated variance of the associated Lagrange multiplier may be too small or even negative, which is a numerical problem. To avoid negative values, tol is the minimum value the diagonal elements can take.

- *summary*: It returns an object of class "summaryGel". It returns the usual tables in estimation methods. The only argument is ... which is passed to the *(*vcov) method.

```
summary(fit)

## Model based on moment conditions
## *********************************
## Moment type: linear
## Covariance matrix: iid
## Number of regressors: 3
## Number of moment conditions: 4
## Number of Endogenous Variables: 1
## Sample size:  50
##
## Estimation:  EL
## Convergence Theta: 0
## Convergence Lambda: 0
## Average |Sum of pt*gt()]|: 4.7261e-17
## |Sum of pt - 1|: 0
##
## coefficients:
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  1.32707    0.86365  1.5366   0.12440
## x1           0.85052    0.11121  7.6480 2.041e-14 ***
## x2          -0.10314    0.06054 -1.7036   0.08845 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Lambdas:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.031751   0.070088  0.4530   0.6505
## x2           0.016730   0.046190  0.3622   0.7172
## z1          -0.027515   0.063545 -0.4330   0.6650
## z2          -0.075503   0.204461 -0.3693   0.7119
##
##  Test E(g)=0
##       Statistics  df   pvalue
## LR:      0.13377   1  0.71455
## LM:      0.13681   1  0.71147
##  J:      0.13212   1  0.71624
```

- specTest It returns the likelihood ratio (LR), Lagrange multiplier (LM) and J test, for the hypothesis $H_0 : \mathrm{E}(g(\theta)) = 0$. The signature is ("gelfit", "missing"), so the second argument, which, is not used. The argument type is set to "ALL" by default, which means that the three test are returned. Alternatively we can specify a single test. It returns an object of class "specTest" which possesses a *print* method.

```
specTest(fit)

##
##  Test E(g)=0
##       Statistics  df   pvalue
## LR:      0.13377   1  0.71455
## LM:      0.13681   1  0.71147
##  J:      0.13212   1  0.71624
```

- *confint*: It returnd confidence itervals for θ or λ By default, a simple Wald type interval is constructed. The parm argument allows to select just a few parameters. By default, it is done for all.

```
confint(fit, parm=1:2, level=0.90)

##
##  Wald type confidence interval
##                   0.05   0.95
## (Intercept)  -0.09351  2.748
## x1            0.66760  1.033

confint(fit, level=0.90, lambda=TRUE)

##
##  Wald confidence interval for Lambda
##                   0.05     0.95
## (Intercept)  -0.08353  0.14704
## x2           -0.05925  0.09271
## z1           -0.13204  0.07701
## z2           -0.41181  0.26081
```

The "vcov" argument is used to provide the method with another covariance matrix. The ... is passed to the *vcov* method.

Confidence interval can also be semi-parametric, by inverting any of the specification test. To undertand how it is constructed, let us define θ_{-i} as the vector θ without θ_i, and $\tilde{\theta}_{-i}(\theta_i)$ its estimate for a given θ_i. Let $S(\theta_i, \theta_{-i})$ be one of the specification test evaluated at $\{\theta_i, \theta_{-i}\}$. Under the null that θ_i is the true value,

$$T(\theta_i) = S(\theta_i, \tilde{\theta}_{-i}(\theta_i)) - S(\hat{\theta}_i, \hat{\theta}_{-i}) \xrightarrow{d} \chi^2_1$$

The following $(1-\alpha)$ confidence interval for θ_i is therefore asymptotically valid (given some regularity conditions):

$$\{\theta_i | T(\theta_i) \leq C_{1-\alpha}\}$$

where $C_{1-\alpha}$ is the $(1-\alpha)$ quantile of the χ^2_1 distribution. To get $\tilde{\theta}_{-i}$, we need to estimate a restricted model for each θ_i and search for $T(\theta_i) = C_{1-\alpha}$. This is done by setting type "invLR", "invLM", or "invJ". A function that returns $[T(\theta_i) - C_{i-\alpha}]$ is passed to the uniroot() function using intervals on both sides of $\hat{\theta}_i$. The argument "fact" is used to control the wideness of the interval passed to uniroot(). By default, the two intervals are (one for the left value and one for the right value of the confidence interval) $[\hat{\theta}_i - 3s_i, \hat{\theta}_i]$ and $[\hat{\theta}_i, \hat{\theta}_i + 3s_i]$, where s_i is the standard error of $\hat{\theta}_i$. If the method fails, it is possible to adjust the intervals with "fact".

```
confint(fit, 1:2, type="invLR", level=0.90)

## Warning:  Could not compute the confidence interval:
##  uniroot failed to find the interval bounds inside  +- fact*SD around the estimate, with
fact= 3 .   Try changing the value of fact

##
##  Confidence interval based on the inversion of the LR test
##                 0.05  0.95
## (Intercept)      NA    NA
## x1           0.6722  1.06
```

A common application in introductory course on EL, is to compute EL confidence interval for a mean. Here, it is done by first creating a model with only one intercept and an intercept as instrument:

```
muModel <- momentModel(x1~1, ~1, data=simData)
```

We then fit the model:

```
muFit <- gelFit(muModel, gelType="EL", tControl=list(method="Brent", lower=-10, upper=10))
muFit@theta

## (Intercept)
##    5.072075
```

The solution is just the sample mean:

```
mean(simData$x1)

## [1] 5.072075
```

We can then compute the EL confidence interval:

```
confint(muFit, type="invLR")

##
##  Confidence interval based on the inversion of the LR test
##              0.025  0.975
## (Intercept)  4.294  5.927
```

We will see below that we can construct GEL intervals in one step, using the *confint* method for "numeric" vectors.

- *update*: The method is used to re-estimate a model with different specifications, keeping all the others equal. Suppose we start with the estimation using the Wu (2005) algorithm:

```
fit <- gelFit(lin)
fit@theta

## (Intercept)          x1          x2
##   1.3270715   0.8505249  -0.1031374
```

We want to re-estimate the model with the same algorithm for λ but a different one for θ:

```
fit2 <- update(fit, coefSlv="nlminb")
fit2@theta

## (Intercept)          x1          x2
##   1.3270608   0.8505255  -0.1031361
```

We can also change the type of GEL

```
fit3 <- update(fit2, gelType="ET", lControl=list())
fit3@theta

## (Intercept)          x1          x2
##   1.3324427   0.8498311  -0.1034840
```

It is also possible to pass a new model to *update*. This is particularly useful to estimate a restricted model with the same numerical methods. This is used by *confint* to compute the interval by the inverse of the specification test.

```
rlin <- restModel(fit3@model, "x1=1")
update(fit3, newModel=rlin)

## Model based on moment conditions
## *********************************
```

```
## Moment type: rlinear
## Covariance matrix: iid
## Number of regressors: 2
## Number of moment conditions: 4
## Number of Endogenous Variables: 1
## Sample size:  50
## Constraints:
##   x1 = 1
## Restricted regression:
##  (y-x1) = (Intercept)+x2
##
## Estimation:  ET
## Convergence Theta:  0
## Convergence Lambda:  0
## coefficients:
## (Intercept)           x2
##  0.23219597  -0.05177569
## lambdas:
## (Intercept)           x2           z1           z2
## -0.23451899   0.02899777   0.16810567  -0.09212368
```

Here, fit3 is an ET estimation. If we restrict the model lin, which is EL, the *update* will also use EL. That's why *restModel* is applied to the model in fit3.

3.2 Robust to misspecification standard errors

The covariance matrix of $\hat{\theta}$ that we described above assumes that the model is correctly specified. In that case, the constraints associated with the moment conditions are asymptotically not binding, which implies that $\hat{\lambda} = O_p(n^{-1/2})$. In misspecified models, some moment conditions are not satisfied and the associated Lagrange multipliers do not converge to 0. Under some regularity conditions (Schennach, 2007), the GEL estimator is still root-n consistent for a pseudo true value. However, inference using the above expression for the covariance matrices is not valid.

One way to obtain valid inference is to rewrite the GEL model into a just-identified GMM one. If we collect the first order conditions of the GEL problem for θ and λ into one vector, we get a GMM model with moment condition $E(\psi_i(\eta)) = 0$, where

$$\psi_i(\eta) = \begin{pmatrix} \rho'(\lambda' g_i(\theta)) G_i(\theta)' \lambda \\ \rho'(\lambda' g_i(\theta)) g_i(\theta) \end{pmatrix},$$

and $\eta = \{\theta', \lambda'\}'$. If we define

$$\hat{\Omega} = \frac{1}{n} \sum_{i=1}^{n} \psi_i(\hat{\eta}) \psi_i(\hat{\eta})'$$

and

$$\hat{\Gamma} = \frac{1}{n} \sum_{i=1}^{n} \nabla_\eta \psi_i(\hat{\eta}),$$

where ∇_η is the gradiant operator, then $\hat{\Gamma}^{-1} \hat{\Omega} \hat{\Gamma}^{-1}$ is a consistent estimator of the variance $\sqrt{n}(\hat{\eta} - \eta^*)$, where η^* is the pseudo true value. The *momFct* method for numeric vector and "gelfit" object computes the $\psi_i(\eta)$ matrix. We could therefore obtain a robust-to-misspecification covariance matrix using the following steps:

```
mod1 <- momentModel(y~x1+x2, ~x2+z1+z2+z3, data=simData)
fit1 <- gelFit(mod1)
```

We create the η vector:

```
eta <- c(coef(fit1), fit1@lambda)
names(eta) <- NULL
```

We create a moment model using the *momFct* method:

```
mod2 <-  momentModel(momFct, fit1, theta0=eta, vcov="MDS")
```

After, we just need to evaluate the model and use the *vcov* method:

```
fit2 <- evalGmm(mod2, eta)
v <- vcov(fit2)[1:3,1:3]
sqrt(diag(v))

##     theta1     theta2     theta3
## 1.08283406 0.14442964 0.06566795
```

Which differs a little from the non-robust standard errors:

```
sqrt(diag(vcov(fit1)$vcov_par))

## (Intercept)          x1          x2
##  0.81563051  0.10697445  0.05761513
```

A quicker way is to use set the option "robToMiss" to TRUE:

```
sqrt(diag(vcov(fit1, robToMiss=TRUE)$vcov_par))

## (Intercept)          x1          x2
##  1.08283406  0.14442964  0.06566795
```

Since the ... of the *summary* method is passed to *vcov*, we can obtain a robust-to-misspecification summary table:

```
summary(fit1, robToMiss=TRUE)@coef

##                Estimate Std. Error   t value     Pr(>|t|)
## (Intercept)  1.20857755 1.08283406  1.116124 2.643689e-01
## x1           0.86256232 0.14442964  5.972197 2.340792e-09
## x2          -0.09612985 0.06566795 -1.463878 1.432274e-01
```

4 The *evalGel* Method

This method create a "gelfit" object without estimation. It is possible to simply fix λ and θ. By default, the GEL type is set to EL:

```
print(evalGel(lin, theta=c(1,2,3), lambda=rep(.1,4)), model=FALSE)

##
## Estimation:  EL
## Convergence Theta:  NA
## Convergence Lambda:  NA
## coefficients:
## (Intercept)           x1           x2
##           1            2            3
## lambdas:
## (Intercept)           x2           z1           z2
##         0.1          0.1          0.1          0.1
```

If the λ is not provided, then it is estimated for the given θ. It is like calling getLambda() with $g(\theta)$ as "gmat". The difference is that a "gelfit" object is returned. It is particularly usefull when we estimate restricted models in which the number of restriction is equal to the number of coefficients. It is use by *confint* method for the confidence interval on the mean (see above):

```
specTest(evalGel(muModel, theta=4), type="LR")

##
##  Test E(g)=0
##       Statistics  df  pvalue
## LR:       7.5096   0      NA
```

The problem with the *evalModel* is that not much effort is put on adjusting the parameters of the model. Above, it says that the degrees of freedom is 0. It is just not an estimation method. A proper estimation of the above would be done with *modelFit*:

```
rmuModel <- restModel(muModel, R=matrix(1,1,1), rhs=4)
specTest(gelFit(rmuModel))

##
##  Test E(g)=0
##       Statistics  df     pvalue
## LR:       7.5096   1  0.0061372
## LM:      10.6270   1  0.0011145
##  J:       5.8924   1  0.0152066
```

In that case *gelFit* is calling *evalGel*.

5 One function to fit them all: gel4

Most users will prefer to avoid going through the steps of building a model and fitting it. The gel4() function build the model, fit it and return the "gelfit" object. The arguments are similar to the ones described above for *modelFit*, *gelSolve* and gelModel().

- Arguments for the model: "g", "x", "theta0", "gelType","rhoFct", "vcov", "grad", "vcovOptions", "centeredVcov", and "data". See gelModel() above for details. The additional arguments "cstLHS" and "cstRHS" are to estimate restricted models. "chstLHS" is passed to *restModel* as "R" and "cstRHS" is passed as "rhs". The default getType is "EL", and the default "vcov" is "MDS"
- Arguments for the estimation: "lamSlv", "coefSlv", "initTheta", "lControl" and "tControl". The default coefSlv is "optim", the default lamSlv is getLambda(), and the default initTheta is "gmm". See *solveGel* and *modelFit* above for more details.
- The argument "getVcov" is passed to *modelFit* as vcov. Therefore, if it is set to TRUE, the covariance matrices are computed.

We can estimate the models from Section 1.1 as follows.

```
fit <- gel4(y~x1+x2, ~x2+z1+z2, data=simData, vcov="iid", gelType="EL",
            lControl=list(algo="Wu"))
print(fit, model=FALSE)

##
## Estimation:  EL
## Convergence Theta:  0
## Convergence Lambda:  0
## coefficients:
## (Intercept)           x1           x2
```

```
##   1.3270715    0.8505249   -0.1031374
## lambdas:
## (Intercept)           x2           z1           z2
##  0.03175148   0.01673016  -0.02751524  -0.07550335
```

To change the initial values in the linear case, we have to provide it and set initTheta to "theta0".

```
fit <- gel4(y~x1+x2, ~x2+z1+z2, data=simData, vcov="iid", gelType="EL",
            lControl=list(algo="Wu"), theta0=c(1,1,0), initTheta="theta0")
coef(fit)

## (Intercept)          x1          x2
##   1.3272620   0.8505003  -0.1031604
```

To estimate the model with the restiction $x1 = x2$, we set $cstLHS$ as R in the *restModel* method:

```
fit <- gel4(y~x1+x2, ~x2+z1+z2, data=simData, vcov="iid", gelType="EL",
            lControl=list(algo="Wu"), cstLHS="x2=x1")
print(fit, model=FALSE)

##
## Estimation:  EL
## Convergence Theta:  0
## Convergence Lambda:  0
## coefficients:
## (Intercept)      (x1+x2)
##    8.835489    -0.313172
## lambdas:
##   (Intercept)             x2             z1             z2
##  0.2705212100  -0.0004247482  -0.2049316654  -0.0190353209
```

For the formula type, a named theta0 is required for the same reason it is required in gelModel(). The names are used to locate the coefficients in the formulas.

```
theta0=c(mu=1,sig=1)
x <- simData$x1
dat <- data.frame(x=x, x2=x^2, x3=x^3, x4=x^4)
gform <- list(x~mu,
              x2~mu^2+sig,
              x3~mu^3+3*mu*sig,
              x4~mu^4+6*mu^2*sig+3*sig^2)
fit <- gel4(g=gform, gelType="EEL", theta0=theta0, vcov="MDS", data=dat)
print(fit, model=FALSE)

##
## Estimation:  EEL
## Convergence Theta:  0
## Convergence Lambda:  0
## coefficients:
##       mu       sig
## 4.708788  4.532422
## lambdas:
##        Mom_1         Mom_2         Mom_3         Mom_4
## -1.068346459   0.427092184  -0.055895513   0.002262411
```

For the function type, theta0 is required because the moment function is evaluated when the model is created. It is also the starting value if initTheta is set to "theta0". the argument x must be provided, because it is the second argument of the moment function

```
fct <- function(theta, x)
   cbind(x-theta[1], (x-theta[1])^2-theta[2],
         (x-theta[1])^3, (x-theta[1])^4-3*theta[2]^2)
dfct <- function(theta, x)
    {
        m1 <- mean(x-theta[1])
        m2 <- mean((x-theta[1])^2)
        m3 <- mean((x-theta[1])^3)
        matrix(c(-1, -2*m1, -3*m2, -4*m3,
                 0, -1, 0, -6*theta[2]), 4, 2)
    }
fit <- gel4(g=fct, x=simData$x3, theta0=c(1,1), grad=dfct, vcov="iid", gelType="ET")
print(fit, model=FALSE)

##
## Estimation:  ET
## Convergence Theta:  0
## Convergence Lambda:  0
## coefficients:
##    theta1     theta2
## 0.1141224  0.8202444
## lambdas:
##          h1           h2           h3           h4
##  0.24454931  -0.79959194  -0.09938698   0.16246548
```

The last type is non-linear models, which requires a named theta0 as for the formula type.

```
theta0=c(b0=1, b1=0, b2=0)
gform <- y~exp(b0+b1*x1+b2*x2)
fit <- gel4(gform, ~z1+z2+z3+x2, theta0=theta0, vcov="MDS", data=simData,
                gelType="HD")
print(fit, model=FALSE)

##
## Estimation:  HD
## Convergence Theta:  0
## Convergence Lambda:  0
## coefficients:
##          b0           b1           b2
##  0.57889968   0.19054605  -0.01927983
## lambdas:
##  (Intercept)            z1            z2            z3            x2
##  0.022346736  -0.014311581  -0.044300403  -0.010552163   0.008644042
```

6 GEL confidence intervals for the mean

The method *confint* can also be used to construct confidence intervals for the mean, or confidence region for two means. In this current version of the package, the default values for the algorithms used to solve for λ and θ cannot be modified. The first method is for "numeric" objects:

```
x2 <- simData$x2
confint(x2, gelType="EL")

##
##  Wald type confidence interval
##     0.025  0.975
## x2  5.209  6.927
```

```
print(confint(x2, gelType="EL", type="invLR"), digits=5)

##
##  Confidence interval based on the inversion of the LR test
##      0.025   0.975
## x2  5.2529  6.9996
```

Which can also be done using the method for "data.frame":

```
confint(simData, "x2", gelType="EL")

##
##  Wald type confidence interval
##     0.025  0.975
## x2  5.209  6.927
```

The arguments "level", "fact" and "type" are as for the method with "gelfit" signature. The "parm" is only used with "data.frame" in which case, it is used to select one or two columns. The argument "vcov" is used to specify the type of data is stored in the object. The options are "iid", the default, or "HAC" for weakly dependent processes.

For "data.frame" and two variables, an object of class "mconfint" is created. The argument "npoints" is the number of points to use to construct the region. The arguments "cores" is for the number of cores to use with mclapply(). For Windows operating systems, it is set to 1. The following shows how to use the *print* and *plot* methods for that class. Here is an example which compares Wald and LR confidence regions.

```
res <- confint(simData, c("x2","y"), npoints=20)
res

## Wald type confidence region
## ***************************
## Level: 0.95
## Number of points: 20
## Variables:
##  Range for mu_x2: [ 5.01, 7.141]
##  Range for mu_y: [3.921, 6.092]
```

We can plot the confidence region using the *plot* method associated with the object created by *confint*:

```
res2 <- confint(simData, c("x2","y"), type="invLR", npoints=20)
c1 <- col2rgb("darkorange2")/255
c1 <- rgb(c1[1],c1[2],c1[3],.5)
c2 <- col2rgb("lightblue2")/255
c2 <- rgb(c2[1],c2[2],c2[3],.5)
plot(res, pch=20, bg=1, Pcol=c1, col=c1, density=10, ylim=c(3.5,6.5),
     xlim=c(4.8,7.5))
plot(res2, pch=20, bg=2, Pcol=c2, col=c2, density=10, add=TRUE)
legend("topright", c("Wald","LR"), col=c(c1,c2), pch=20 , bty="n")
```

The arguments "main", "xlab" "ylab", "pch", "ylim", "xlim", "Pcol" (the color of the points) and "bg" are passed to plot.default(), and all other arguments are passed to polygon().

References

S. Anatolyev. Gmm, gel, serial correlation, and asymptotic bias. *Econometrica*, 73:983–1002, 2005.

B. Antoine, H. Bonnal, and E. Renault. On the efficient use of the informational content of estimating equations: Implied probabilities and euclidean empirical likelihood. *Journal of Econometrics*, 138: 461–487, 2007.

Pierre Chaussé. Computing generalized method of moments and generalized empirical likelihood with r. *Journal of Statistical Software*, 34(11):1–35, 2010. URL `https://www.jstatsoft.org/v34/i11/`.

Pierre Chaussé. *gmm: Generalized Method of Moments and Generalized Empirical Likelihood*, 2019. R package version 1.6-3.

L. P. Hansen, J. Heaton, and A. Yaron. Finit-sample properties of some alternative gmm estimators. *Journal of Business and Economic Statistics*, 14:262–280, 1996.

Y. Kitamura and M. Stutzer. An information-theoretic alternative to generalized method of moments estimation. *Econometrica*, 65(5):861–874, 1997.

Philip Leifeld. texreg: Conversion of statistical model output in R to LaTeX and HTML tables. *Journal of Statistical Software*, 55(8):1–24, 2013. URL `https://www.jstatsoft.org/v55/i08/`.

W. K. Newey and R. J. Smith. Higher order properties of gmm and generalized empirical likelihood estimators. *Econometrica*, 72:219–255, 2004.

A. B. Owen. *Empirical Likelihood*. Chapman and Hall, 2001.

S. M. Schennach. Point estimation with exponentially tilted empirical likelihood. *Econometrica*, 35(2): 634–672, 2007.

R. J. Smith. Gel criteria for moment condition models. *Econometric Theory*, 27(6):1192–1235, 2011.

C. Wu. Algorithms and R codes for the pseudo empirical likelihood method in survey sampling. *Survey Methodology*, 31(2):239, 2005.

A Some extra codes

The following *extract* is used with the "texreg" package of Leifeld (2013) to produce nice latex tables.

```
library(texreg)
setMethod("extract", "gelfit",
          function(model, includeSpecTest=TRUE,
                   specTest=c("LR","LM","J"), include.nobs=TRUE,
                   include.obj.fcn=TRUE, ...)
              {
                  specTest <- match.arg(specTest)
                  s <- summary(model, ...)
                  wspecTest <- grep(specTest, rownames(s@specTest@test))
                  spec <- modelDims(model@model)
                  coefs <- s@coef
                  names <- rownames(coefs)
                  coef <- coefs[, 1]
                  se <- coefs[, 2]
                  pval <- coefs[, 4]
                  n <- model@model@n
                  gof <- numeric()
                  gof.names <- character()
                  gof.decimal <- logical()
                  if (includeSpecTest) {
                      if (spec$k == spec$q)
```

	Model 1	Model 2
(Intercept)	−1.0507	−0.4054
	(0.6155)	(0.9789)
x1	1.1968***	1.1129***
	(0.1254)	(0.1505)
x2		−0.0355
		(0.0504)
LR-test Statistics	0.5866	0.1901
LR-test p-value	0.7458	0.6628
Num. obs.	50	50

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 1: Statistical models

```
                    {
                        obj.fcn <- NA
                        obj.pv <- NA
                    } else {
                        obj.fcn <- s@specTest@test[wspecTest,1]
                        obj.pv <- s@specTest@test[wspecTest,3]
                    }
                gof <- c(gof, obj.fcn, obj.pv)
                gof.names <- c(gof.names,
                               paste(specTest,"-test Statistics", sep=""),
                               paste(specTest,"-test p-value", sep=""))
                gof.decimal <- c(gof.decimal, TRUE, TRUE)
            }
            if (include.nobs == TRUE) {
                gof <- c(gof, n)
                gof.names <- c(gof.names, "Num.\\ obs.")
                gof.decimal <- c(gof.decimal, FALSE)
            }
            tr <- createTexreg(coef.names = names, coef = coef, se = se,
                               pvalues = pval, gof.names = gof.names, gof = gof,
                               gof.decimal = gof.decimal)
            return(tr)
        })
```

Here is an example

```
fit1 <- gel4(y~x1, ~x2+x3+x4, data=simData)
fit2 <- gel4(y~x1+x2, ~x2+x3+x4, data=simData)
texreg(list(fit1,fit2), digits=4)
```