Linearized GMM Estimator

Doosoo Kim

Department of Economics, Toronto Metropolitan University

0. Summary

- Nonlinear generalized method of moments (GMM) estimators often encounter computational issues when the moment conditions are **over-identifying and nonlinear**.
- I propose a novel GMM estimator based on **linearized** moment conditions approximated around an underlying **exactly-identified** (or over-identified) parameter estimate.
- For any given standard moment condition, I prove the existence of such an underlying parameter, and introduce a straightforward algorithm for its identification. The added dimensions in the underlying exactly-identified parameter can be estimated one element at a time, separately.
- This estimator exhibits improved computational properties while maintaining first-order asymptotic efficiency.
- The enhancement arises from (i) the better-behaved curvature of the GMM objective function (e.g. strict local convexity) for estimating the underlying parameter, and (ii) the availability of a closed-form solution for the final estimate.
- The method is applied to Ahn, Lee, and Schumidt's (2013) panel data model with multiple time-varying individual effects.

1. Introduction

- Supose a standard moment condition

$$E\left[g_i\left(\beta_o\right)\right] = 0$$

where
$\{w_i\}_{i=1}^N$; i.i.d data,
$g_i(\beta) = g(w_i, \beta)$; q-dimensional moment function
β_o ; p-dimensional parameter
$q \geq p$; order condition

- Q: Is it possible to find $g_i(\gamma(\beta))$ such that

$$g_i\left(\gamma\left(\beta\right)\right) = g_i\left(\beta\right) \;\; \forall \beta \in \Theta_\beta$$

with exactly-identified γ_o and full rank $E[\partial g(\gamma_o)/\partial\gamma']$?

- If it is, we can linearly expand the moment function around $\hat{\gamma}$

$$g_i\left(\gamma\left(\beta\right)\right) \approx g_i\left(\hat{\gamma}\right) + \frac{\partial g_i\left(\hat{\gamma}\right)}{\partial \gamma'}\left(\gamma\left(\beta\right) - \hat{\gamma}\right)$$

and, noting that $\sum_{i=1}^N g_i(\hat{\gamma}) = 0$, solve

$$\min_{\beta}\left(\gamma\left(\beta\right) - \hat{\gamma}\right)' G_N\left(\hat{\gamma}\right)' \hat{W} G_N\left(\hat{\gamma}\right)\left(\gamma\left(\beta\right) - \hat{\gamma}\right)$$

where $G_N(\gamma) \equiv \frac{1}{N}\sum \frac{\partial g_i(\gamma)}{\partial \gamma'}$.

  - The same problem is solved by the classicle minimum distance estimator given $G_N(\hat{\gamma})' \hat{W} G_N(\hat{\gamma})$ (Kim, 2020).

- Given the existence of $g_i(\gamma(\beta))$, the approach based on the moment function approximated to a generic degree (e.g. second-order, third-order) is also interesting. The linearized one has useful implications on numerical issues in GMM estimation. Two benefits are discussed in the following.

2. A Cause of Computational Issues

- We focus on the curvature of the GMM objective function used in the **second step** of Hansen's two-step GMM estimator. (e.g. Ahn, Lee, and Schmidt, 2013)
- Two-step GMM
  - Step1: a preliminary estimate $\tilde{\beta}$ is computed, for example with the identity weighting matrix.
  - Step2: the efficient estimator $\hat{\beta}_{TS}$ minimizes

$$Q_N\left(\beta\right) \equiv g_N\left(\beta\right)' \hat{W} g_N\left(\beta\right)$$

  where $g_N(\beta) \equiv \frac{1}{N}\sum_{i=1}^N g_i(\beta)$ and $\hat{W} \equiv (\frac{1}{N}\sum_{i=1}^N g_i(\tilde{\beta}) g_i(\tilde{\beta})')^{-1}$.

- The second order derivative of $Q_N(\beta)$

$$\frac{1}{2}\frac{\partial}{\partial\beta'}\left(\frac{\partial Q_N\left(\beta\right)}{\partial\beta}\right) = \underbrace{G_N\left(\beta\right)' \hat{W} G_N\left(\beta\right)}_{\text{pd if } \hat{W} \text{ pd and } G_N(\beta) \text{ full rank}} + \underbrace{\left(I_p \otimes g_N\left(\beta\right)' \hat{W}\right)\frac{\partial}{\partial\beta'} vec\left[G_N\left(\beta\right)\right]}_{\text{not necessarily pd. } =0 \text{ if linear or exact-id}}$$

where $G_N(\beta) = \frac{1}{N}\sum_{i=1}^N \frac{\partial g_i(\beta)}{\partial \beta'}$.

- **Benefit 1:** the curvature near its optimal point $\hat{\gamma}$ would be better-behaved. Also, the optimal value can be known.
  - When the sample moment condition is solvable, we can use numerical algorithms for solving nonlinear equation systems (e.g. Broyden-Powell method).
  - When it is not solvable, we would still have $\left\|\hat{W}^{1/2} g_N\left(\hat{\gamma}\right)\right\| \leq \left\|\hat{W}^{1/2} g_N(\tilde{\beta})\right\|$ by construction of γ so that the influence of the second term is reduced.
- **Benefit 2:** A closed form expression for the estimator of β would be available given the γ estimate.

3. Existence of An Underlying Parameter: Reparametrization

Under standard assumptions, there exists reparametrization such that (a) $g_i(\gamma(\beta)) = g(\beta) \; \forall \beta \in \Theta_\beta$ (b) $E[g(\gamma)] = 0$ if and only if $\gamma = \gamma_o$ and (c) $rank\left(E[\partial g(\gamma_o)/\partial\gamma']\right) \geq p$. If reparametrization can depend on the value of $E[\partial g_i(\beta_o)/\partial\beta]$, it is possible to make $E[\partial g(\gamma_o)/\partial\gamma']$ full rank.

4. Toy Example

For a scalar parameter β, consider $q > 1$ moment condition

$$E\left[g_{i1}\left(\beta\right)\right] = 0$$
$$\vdots$$
$$E\left[g_{iq}\left(\beta\right)\right] = 0$$

The key idea is "reparametrization":

$$E\left[g_{i1}\left(\gamma_1\right)\right] = 0$$
$$\vdots$$
$$E\left[g_{iq}\left(\gamma_q\right)\right] = 0$$

where

$$\gamma\left(\beta\right) = \left[\beta \cdots \beta\right]'$$

5. Methods

- Construction of $g(\gamma(\beta))$ (based on the proof idea)
  - Denote $\gamma = (\gamma_1', \gamma_2')'$ where γ_1 is p-dim and γ_2 is $(q-p)$-dim.
  - First, consider trivial reparametrization: put $\beta = \gamma_1$ for all moments.
  - Second, we select $q-p$ moments and pick one excess parameter per moment to define γ_2.
  - We assume $E\left[\frac{\partial g_i(\beta_o)}{\partial\beta}\right]$ has full rank. WLOG, Stochastic Gaussian Elimination will find an invertible A_0 such that

$$E\left[A_0\frac{\partial g_i\left(\beta_o\right)}{\partial\beta}\right] = \begin{bmatrix} G_1 \\ 0_{(q-p)\times p} \end{bmatrix}$$

  where G_1 is an invertible upper triangular matrix. For simplicity, let's assume G_1 is an identity matrix, I_p
  - Define $A = RA_0$ where R adds the first row of G_1 to all and each of the lower zero matrix $0_{(q-p)\times p}$.
  - Reparametrize the first element of β to be an element in γ_2 in the lower $(q-p)$ rows of $Ag_i(\beta)$. Then, the Jacobian with respect to γ will look like

$$\begin{bmatrix} G_1 & 0 \\ 0 & I_{(q-p)} \end{bmatrix}$$

- Efficient LGMM
  - Given $\hat{g}_i(\gamma(\beta))$, estimate γ by solving

$$\min\left(\frac{1}{N}\sum \hat{g}_i\left(\gamma\right)\right)'\left(\frac{1}{N}\sum \hat{g}_i\left(\gamma\right)\right)$$

    - Each of excess parameters γ_2 can be separately estimated given γ_1 estimate.
  - closed form solution for $\hat{\beta}$.
    - LGMM

$$\hat{\beta}_{LGMM} = \left[\hat{\Gamma}'\hat{G}_{N,\gamma}'\hat{\Omega}^{-1}\hat{G}_{N,\gamma}\hat{\Gamma}\right]^{-1}\left[\hat{\Gamma}'\hat{G}_{N,\gamma}'\hat{\Omega}^{-1}\hat{G}_{N,\gamma}\hat{\gamma} - \Gamma'\hat{G}_{N,\gamma}'\hat{\Omega}^{-1}\hat{g}_N\left(\hat{\gamma}\right)\right]$$

    - LGMM-LITE (with $\gamma(\beta) = \beta$; the preliminary estimate is taken as $\hat{\gamma}$)

$$\hat{\beta}_{LGMM-LITE} = \tilde{\beta} - \left[\tilde{G}_{N,\beta}'\tilde{\Omega}^{-1}\tilde{G}_{N,\beta}\right]^{-1}\tilde{G}_{N,\beta}'\tilde{\Omega}^{-1}g_N(\tilde{\beta})$$

- The resulting estimator maintains the first-order asymptotic efficiency. Asymptotic properties of the estimator and related discussions can be found in the draft.

6. Simulation

- Ahn, Lee, and Schmidt (2013). Two slope parameters.

	Draws	N	T	Bias1	RMSE1	Bias2	RMSE2	Failure
2step	1000	100	10	.245	.554	.242	.546	59
L-Lite	1000	100	10	.230	.538	.227	.530	0
LGMM	1000	100	10	.252	.512	.210	.566	0

7. Further Questions

- Higher-order approximations?
- Without estimating the Jacobian matrix?

Contact Information

- Email: doosoonet@gmail.com
- The paper and STATA package are available at http://doosoo.kim