Vol. 36 (2016)
No. 2

数 学 杂 志
J. of Math. (PRC)

THE LAW OF ITERATED LOGARITHM FOR GMM ESTIMATORS

YANG Yan-qiu

(*School of Mathematics and Statistics, Wuhan University, Wuhan 430072*)

Abstract: In this article, we study the problem on GMM estimators. By using the strong limit condition, we obtain the result that the generalized method of moments obeys the law of iterated logarithm.

Keywords: generalized method of moments; strong consistency; law of iterated logarithm

2010 MR Subject Classification: 62F12; 62N02

Document code: A **Article ID:** 0255-7797(2016)02-0246-07

1 Introduction

Using estimating equations to estimate unknown parameters was widely studied both in empirical applications and theoretical studies, see Hall [3] for reviews of this method. If the number of estimating equations is identical to the number of unknown parameters (just-identification), we can use usual moments method to estimate the unknown parameters and further study its statistical asymptotic behavior. But in applications, especially in econometrics time series datas and longitudinal datas analysis, it is often the case that the number of estimating equations is larger than the number of unknown parameters (called over-identification) and in this case the solution of moments estimating equations does not exists in general. To deal with this problem, several estimation methods were proposed in the literature, the most popular were the generalized method of moments (GMM) by Hansen in [4] and generalized empirical likelihood (GEL) by Smith in [8]. This paper mainly focused on the GMM estimators.

To fix the main idea of GMM, suppose we have observed a random sample $(X_1, \cdots, X_n)$ from some population X with unknown parameters $\theta \in \Theta$. The true value θ_0 satisfies the population moment conditions

$$\mathbf{E}[g(X, \theta_0)] = 0, \tag{1.1}$$

where Θ is a compact such set of $\mathbb{R}^p, p \in \mathbb{N}_+$ and g is a $\mathbb{R}^q, q \in \mathbb{N}_+$ valued measurable function. In the over identified case, i.e., $q > p$, Hansen (see [4]) introduced the GMM

* **Received date:** 2015-03-07 **Accepted date:** 2015-04-21

Foundation item: Supported by National Natural Science Foundation of China (11171262).

Biography: Yang Yanqiu(1990–), female, born at Liaocheng, Shandong, master, major in probability.

estimator of θ_0 as

$$\hat{\theta}_{1n} = \arg\min_{\theta\in\Theta} \left(\frac{1}{n}\sum_{i=1}^{n} g(X_i,\theta)\right)' W_n \left(\frac{1}{n}\sum_{i=1}^{n} g(X_i,\theta)\right), \tag{1.2}$$

where $'$ means transposition operation of matrix and W_n is a positive semi-definite weighting matrix that converges in probability to a positive definite constants matrix W. If $W_n \equiv I$, then the $\hat{\theta}_{1n}$ is the point of θ so that minimize the sample covariance matrix. It is known that if the model is not misspecified and under some mild regularity conditions, Hansen (see [4]) established the strong consistency and asymptotic normality of the GMM estimator. Precisely, he proved

$$\sqrt{n}(\hat{\theta}_{1n} - \theta_n) \to_d N(0, V_W), \quad V_W = (G'WG)^{-1}G'W\Sigma WG(G'WG)^{-1}, \tag{1.3}$$

where $G = \mathbf{E}(\nabla g_\theta(X,\theta_0)), \Sigma = \mathbf{E}[g(X,\theta_0)g'(X,\theta_0)]$. Observing that the asymptotic variance V_W depends on W. When constructing the confidence intervals of θ_0, we hope V_W as small as possible (in the positive semi-definite sense). It is easy to find that V_W is minimized at $V_W = \Sigma^{-1}$. Since Σ is unknown, we can use $\hat{\theta}_{1n}$ as a preliminary estimator, and naturally defined

$$\hat{\Sigma}_n^{-1} = \left(\frac{1}{n}\sum_{i=1}^{n} g(X_i,\hat{\theta}_{1n})g(X_i,\hat{\theta}_{1n})'\right)^{-1}$$

as a estimator of Σ^{-1}. Then we drive a two-step GMM estimator:

$$\hat{\theta}_{2n} = \arg\min_{\theta\in\Theta} \left(\frac{1}{n}\sum_{i=1}^{n} g(X_i,\theta)\right)' \hat{\Sigma}_n^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} g(X_i,\theta)\right), \tag{1.4}$$

and for convenience we call $\hat{\theta}_{1n}$ as one-step GMM estimator.

The large sample behaviors of one-step and two-step has been studied extensively. Under mild conditions, $\hat{\Sigma}_n^{-1}$ converges weakly to Σ^{-1} and $\hat{\theta}_{2n}$ is consistent and asymptotic normality (see [2]). Otsu studied the large deviation principle of the these estimators in [5] and the moderate deviation principle was also obtained in [6]. This paper is devoted to study the law of iterated logarithm for GMM estimators.

The paper is organized as follows: In Section 2, we state the basic assumptions and main results of the paper, the proof of the main results are given in Section 3.

2 Main Results

We first give some notations. For any $m \times n$ matrix $A = (a_{ij})_{m\times n}$, we write A' as its transposition, and define its norm as

$$\|A\| = \sqrt{\mathrm{tr}(AA')} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}.$$

Denote

$$G(X,\theta)=\nabla g_\theta(X,\theta),\qquad G=\mathbf{E}\left(\nabla g_\theta(X,\theta_0)\right),$$
$$\hat g_n(\theta)=\frac{1}{n}\sum_{i=1}^n g(X_i,\theta),\qquad \hat G_n(\theta)=\frac{1}{n}\sum_{i=1}^n G(X_i,\theta),$$
$$Q(\theta_0)=(q_{ij}(\theta_0))_{p\times p}=G'WG$$

and $h_i(\theta_0)=(h_{ij}(\theta_0))_{p\times 1}:=G'Wg(X_i,\theta_0)$, in the following we assume W is positive definite constants matrix, so $Q(\theta_0)$ is positive, we can define $A(\theta_0)=(a_{ij}(\theta_0))_{p\times p}=-Q(\theta_0)^{-1}$, which is a negative definite matrix.

Following are the assumptions of the paper.

Assumption 2.1 (1) Θ is compact and $\theta_0\in \mathrm{int}(\Theta)$. There exists a measurable function $L:\mathbb{R}\to[0,\infty)$ so that for a.e. x, $\|g(x,\theta_1)-g(x,\theta_2)\|\le L(x)\|\theta_1-\theta_2\|,\forall\,\theta_1,\theta_2\in\Theta$. For some $\delta_1>0$ and all $\theta\in\Theta$, $\mathbf{E}\|g(X,\theta)\|^{2+\delta_1}<\infty$ and $\mathbf{E}L(X)<\infty$.

(2) There exists a measurable function $H:\mathbb{R}\to[0,\infty)$ and $\delta_2>0$ such that $\|G(x,\theta)-G(x,\theta_0)\|\le H(x)\|\theta-\theta_0\|$ holds for a.e. x and $\theta\in U(\theta_0,\delta_2):=\{\theta:\|\theta-\theta_0\|<\delta_2\}$. Besides, we assume $\mathbf{E}H(X)<\infty$, $\mathbf{E}H(X)^2<\infty$, $\mathbf{E}\|G(X,\theta)\|<\infty$ and $\mathbf{E}\|G(X,\theta)\|^2<\infty$, for all $\theta\in\Theta$. The $(q\times p)$matrix G has the full column rank and Σ is positive definite.

For the weighting matrix W_n, we assume

Assumption 2.2 $\{W_n\}_{n\ge 1}$ is a sequence of positive semi-definite matrices, and W_n converges weakly to W in the sense of matrix norm, where W is a positive definite matrix of constants. Besides, $\sup_n \mathbf{E}\|W_n\|<\infty$.

Suppose that Assumptions 2.1, 2.2 hold, our main results are

Theorem 2.1 Let $\hat\theta_{1n}=(\hat\theta_{11},\hat\theta_{21},\cdots,\hat\theta_{p1})'$ and $\hat\theta_{2n}=(\hat\theta_{12},\hat\theta_{22},\cdots,\hat\theta_{p2})'$ be the one-step and two-step GMM estimators, then for any $\theta_0=(\theta_{10},\theta_{20}\cdots,\theta_{p0})'\in\Theta$, we have

$$\limsup_{n\to\infty}\frac{n}{\sqrt{2n\log\log n}}|\hat\theta_{si}-\theta_{s0}|=\varphi(s), s=1,2,\cdots,p, i=1,2, \tag{2.1}$$

where

$$\varphi^2(s)=\sum_{m=1}^p\sum_{k=1}^p a_{sm}(\theta_0)a_{sk}(\theta_0)\mathbf{E}\left(h_{im}(\theta_0)h_{ik}(\theta_0)\right), s=1,2,\cdots,p.$$

Remark 2.1 Hansen (see [4]) established the strong consistency of the GMM estimator. Under Assumptions 2.1, 2.2 and classical strong law of large numbers we have

$$P\left(\lim_{\delta\to 0}\limsup_{\substack{n\to\infty\\ \theta\in U(\theta_0,\delta)}}\|\hat G_n(\theta)-\hat G_n(\theta_0)\|=0\right)=1,$$
$$\|\hat G_n(\theta_0)-G\|\to 0,\qquad \hat G_n(\hat\theta_{1n})'W\hat g_n(\theta_0)\to 0\quad \text{a.s. } n\to\infty.$$

Because the independence is necessary in proof of the next lemma and theorem,we first suppose $W_n=W$. Then under the strong limit condition of W_n and W, we get the results for W_n.

3 Law of Iterated Logarithm

3.1 The Lemma and its Proof

Lemma 3.1 (see [7], Theorem 4, Chap. X) Let $\{U_i : i \geq 1\}$ be a sequence of independent random variables with $\mathbf{E}U_i = 0,\ i \geq 1, \mathbf{E}U_i^2 < \infty$. Setting $S_n^2 = \sum\limits_{i=1}^{n} \mathbf{E}U_i^2$, if $\{U_i, i \geq 1\}$ satisfies the conditions

$$\liminf_{n\to\infty} \frac{S_n^2}{n} > 0, \tag{3.1}$$

and for some $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbf{E}[U_k^2 |\log |U_k||^{1+\delta}] < \infty. \tag{3.2}$$

Then

$$\limsup_{n\to\infty} \frac{|\sum\limits_{i=1}^{n} U_i|}{\sqrt{2S_n^2 \log\log S_n^2}} = 1 \qquad \text{a.s..} \tag{3.3}$$

Remark 3.1 The following easy verified condition is sufficient for (3.2),

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbf{E}(|U_k|^{2+\delta}) < \infty.$$

Since for x large enough, we have elementary inequality $(\log x)^{1+\delta} < x^\delta$ for any $\delta > 0$.

Lemma 3.2 We denote

$$\omega_i(s) \quad = \quad \sum_{m=1}^{p} a_{sm}(\theta_0) h_{im}(\theta_0)$$

and $S_n^2 = \sum\limits_{i=1}^{n} \mathbf{E}\omega_i^2(s)$. Then

$$\limsup_{n\to\infty} \frac{|\sum\limits_{i=1}^{n} \omega_i(s)|}{\sqrt{2S_n^2 \log\log S_n^2}} = 1 \qquad \text{a.s..} \tag{3.4}$$

Proof Observe that $\mathbf{E}(h_{im}(\theta_0)) = G'W\mathbf{E}[g(X_i, \theta_0)] = 0$ and $\{h_{im}, i \geq 1\}$ is a sequence of independent random variables. Consequently,

$$\mathbf{E}\omega_i(s) = \sum_{m=1}^{p} a_{sm}(\theta_0) \mathbf{E}(h_{im}(\theta_0)) = 0.$$

By virtue of Assumptions 2.1, 2.2, we have

$$\sup_{i\geq 1} \max_{1\leq m\leq p} \mathbf{E}\|h_{im}\|^{2+\delta} \leq \sup_{i\geq 1} \mathbf{E}\|GWg(X_i, \theta_0)\|^{2+\delta} < \infty.$$

Therefore

$$\sup_{i\geq 1}\mathbf{E}|\omega_i(s)|^{2+\delta}<\infty, \qquad \limsup_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}|\omega_i(s)|^{2+\delta}<\infty.$$

On the other hand, it is obviously that $\lim\limits_{n\to\infty}\frac{S_n^2}{n}=\mathbf{E}\omega_1^2(s)>0$, consequently, it follows from Lemma 3.1 that (3.4) holds.

Now it ready to give the proof of Theorem 2.1.

3.2 The Proof of Theorem 2.1

For the consistence between theorem and lemma, we set first

$$\theta_n(t)=(1-t)\theta_0+t\hat{\theta}_{1n}, \qquad u_n(t)=\hat{G}_n(\hat{\theta}_{1n})'W\hat{g}_n(\theta_n(t)), \qquad 0\leq t\leq 1. \tag{3.5}$$

Hence $\theta_n(1)=\hat{\theta}_{1n}$, $u_n(1)=\hat{G}_n(\hat{\theta}_{1n})'W\hat{g}_n(\hat{\theta}_{1n})=0$. From (3.5), we have

$$u_n(0)=u_n(1)-\int_0^1 u_n'(t)dt=-\int_0^1 Q_n(\theta_n(t))(\hat{\theta}_{1n}-\theta_0)dt, \tag{3.6}$$

where

$$Q_n(\theta_n(t))=\hat{G}_n(\hat{\theta}_{1n})'W\hat{G}_n(\theta_n(t)).$$

We denote

$$\varphi^2(s)=\lim_{n\to\infty}\frac{S_n^2}{n}=\sum_{m=1}^{p}\sum_{k=1}^{p}a_{sm}(\theta_0)a_{sk}(\theta_0)\mathbf{E}\left(h_{im}(\theta_0)h_{ik}(\theta_0)\right). \tag{3.7}$$

It follows from Remark 2.1 and (3.5) that $\|\theta_n(t)-\theta_0\|\to 0$ a.s. for all $t\in(0,1)$. By Remark 2.1 and Assumptions 2.1, 2.2, we have

$$\begin{aligned}
&\limsup_{n\to\infty}\|\hat{G}_n(\hat{\theta}_{1n})'W\hat{G}_n(\theta_n(t))-G'WG\|\\
\leq&\limsup_{n\to\infty}\Big\{\|(\hat{G}_n(\hat{\theta}_{1n})-\hat{G}_n(\theta_0))'W\hat{G}_n(\theta_n(t))\|+\|(\hat{G}_n(\theta_0)-G)'W\hat{G}_n(\theta_n(t))\|\\
&+\|(G'W(\hat{G}_n(\theta_n(t))-\hat{G}_n(\theta_0))\|+\|G'W(\hat{G}_n(\theta_0)-G)\|\Big\}\to 0,
\end{aligned}$$

almost surely. It follows immediately from above that

$$\lim_{n\to\infty}\int_0^1 Q_n(\theta_n(t))dt=\int_0^1\lim_{n\to\infty}Q_n(\theta_n(t))dt=\int_0^1 Q(\theta_0)dt=Q(\theta_0). \tag{3.8}$$

Denote

$$\Delta_n:=Q^{-1}(\theta_0)-\left(\int_0^1 Q_n(\theta_n(t))dt\right)^{-1}. \tag{3.9}$$

Since $Q(\theta_0)$ is positive definite, thus $\|Q^{-1}(\theta_0)\|<\infty$. And it follows from (3.8) that there exists $N>0$ so that

$$\sup_{n>N}\left\|\left(\int_0^1 Q_n(\theta_n(t))dt\right)^{-1}\right\|<\infty.$$

Consequently,

$$\|\Delta_n\| \leq \|Q^{-1}(\theta_0)\|\|Q(\theta_0) - \int_0^1 Q_n(\theta_n(t))dt\|\|(\int_0^1 Q_n(\theta_n(t))dt)^{-1}\|.$$

When $\|\hat{\theta}_{1n} - \theta_0\| \to 0, n \to \infty$, we get $\Delta_n \to 0$. From Remark 2.1 and Assumptions 2.1, 2.2, it follows that

$$\|u_n(0) - G'W\hat{g}_n(\theta_0)\| = \|(\hat{G}_n(\hat{\theta}_{1n}) - \hat{G}_n(\theta_0))'W\hat{g}_n(\theta_0) + (\hat{G}_n(\theta_0) - G)'W\hat{g}_n(\theta_0)\| \to 0$$

almost surely. Next, from (3.6) and (3.9), we have

$$\begin{aligned}\frac{1}{\sqrt{2n\log\log n}}(\hat{\theta}_{1n} - \theta_0) &= \frac{1}{\sqrt{2n\log\log n}}\left(-\int_0^1 Q_n(\theta_n(t))dt\right)^{-1} u_n(0)\\ &= \frac{1}{\sqrt{2n\log\log n}}(\Delta_n - Q^{-1}(\theta_0))u_n(0)\\ &= -\frac{1+o(1)}{\sqrt{2n\log\log n}}Q^{-1}(\theta_0)G'W\hat{g}_n(\theta_0).\end{aligned}$$

Thus for any $s = 1, \cdots, p$, by virtue of Lemma 3.2, we finally have

$$\begin{aligned}&\limsup_{n\to\infty} \frac{1}{\sqrt{2n\log\log n}} n|\hat{\theta}_{s1} - \theta_{s0}|\\ =&\limsup_{n\to\infty} \frac{1}{\sqrt{2n\log\log n}} n\left|(a_{s1}(\theta_0), \cdots, a_{sp}(\theta_0))\left(\frac{1}{n}\sum_{i=1}^n G'Wg(X_i,\theta_0)\right)\right|\\ =&\limsup_{n\to\infty} \frac{1}{\sqrt{2n\log\log n}}\left|\sum_{i=1}^n\sum_{m=1}^p a_{sm}(\theta_0)h_{im}(\theta_0)\right|\\ =&\limsup_{n\to\infty}\sqrt{\frac{S_n^2\log\log S_n^2}{n\log\log n}}\sqrt{\frac{1}{2S_n^2\log\log S_n^2}}\left|\sum_{i=1}^n \omega_i(s)\right| = \varphi(s).\end{aligned}$$

This complete the whole proof.

Now we consider the second step GMM estimator $\hat{\theta}_{2n}$ in (1.4). We suppose the new assumption

Assumption 3.1 There exists $\delta_3 > 0$, such that $\mathbf{E}\|g(X,\theta_0)g(X,\theta_0)'\|^{2+\delta_3} < \infty$, and $\mathbf{E}\|L(X)^2\|^{2+\delta_3} < \infty$.

It follows from Assumption 3.1 and strong law of large numbers that

$$\hat{\Sigma}_n \xrightarrow{\text{a.s.}} \mathbf{E}(g(X,\theta_0)g(X,\theta_0)').$$

So we have $W_n = (\hat{\Sigma}_n)^{-1} \xrightarrow{\text{a.s.}} \Sigma^{-1} = W$. We replace $\hat{\theta}_{1n}, W$ with $\hat{\theta}_{2n}, W_n$. For $\hat{\theta}_{2n} = (\hat{\theta}_{12}, \hat{\theta}_{22}, \cdots, \hat{\theta}_{p2})'$, (2.1) holds.

References

[1] Fan Jun, Gao Fuqing. Moderate deviations and the law of iterated logarithm for LSE in linear model[J]. J. Math., 2007, 27(1): 60–64.

[2] Hall A R. Generalized method of moments[M]. UK: Oxford University Press, 2005.

[3] Hall A R, Inoue A. The large sample behaviour of the generalized method of moments estimator in misspecified models[J]. J. Econ., 2003, 114: 361–394.

[4] Hansen L P. Large sample properties of generalized method of moments estimators[J]. J. Econ., 1982, 50: 1029–1054.

[5] Otsu T. Large deviations of generalized method of moments and empirical likelihood estimators[J]. J. Econ., 2011, 14: 321–329.

[6] Otsu T. Moderate deviations of generalized method of moments and empirical likelihood estimators[J]. J. Multi. Anal., 2011, 102: 1203–1216.

[7] Petrov V V. Sums of independent random variables[M]. Translated by Brown A A. Berlin: Springer Verlag, 1975.

[8] Smith R J. Alternative semi-parametric likelihood approaches to generalized method of moments estimation[J]. J. Econ., 1997, 107:503–519.

[9] Zhu Qiang, Gao Fuqing. A Chung type law of the iterated logarithm of MLE in random censoring model with incomplete information[J]. Acta Math. Scientia, 2007, 27(A(4)): 672–681.

广义矩估计的重对数律

杨艳秋

(武汉大学数学与统计学院, 湖北 武汉 430072)

摘要: 本文研究了广义矩估计的性质. 利用强相合性的条件, 得到了广义矩估计满足重对数律的结果.

关键词: 广义矩估计; 强相合性; 重对数律

MR(2010)**主题分类号**: 62F12; 62N02 **中图分类号**: O211