Bayesian Analysis in Moment Inequality Models
Supplement Material

Yuan Liao Wenxin Jiang

November 14, 2008

Abstract

This is a supplement material of Liao and Jiang (2008). We consider the case when the identified region has no interior. We show that the posterior distribution converges to zero on any δ-contraction outside the identified region exponentially fast, and is bounded below by a polynomial rate on any neighborhood of element in a dense subset, defined by both exact moment conditions and strict moment inequalities. Hence the consistent estimation of the identified region can be constructed based on the log-posterior pdf.

1 Case When $int(\Omega)$ is Empty

When Ω has no interior, moment inequality models may contain exact moment conditions.

$$\begin{aligned} & Em_{1j}(X,\theta_0) \geq 0, j = 1,...,r \\ & Em_{2j}(X,\theta_0) = 0, j = 1,...,p \end{aligned} \tag{1.1}$$

Moon and Schorfheide (2006) have considered the estimation problem assuming θ_0 is point identified by the exact moment conditions. Let

$$m_1(X,\theta) = (m_{11}(X,\theta),...,m_{1r}(X,\theta))^T, m_2(X,\theta) = (m_{21}(X,\theta),...,m_{2p}(X,\theta))^T$$

If $p \geq \dim(\theta_0)$, and there doesn't exist a pair of moment functions (m_{2i}, m_{2j}) such that $\{\theta \in \Theta : Em_{2i}(X,\theta) = 0\} = \{\theta : Em_{2j}(X,\theta) = 0\}$, then θ_0 is point identified by $Em_2(X,\theta_0) = 0$. Moon and Schorfheide (2006) show that by using the overidentifying information provided by $Em_1(X,\theta_0) \geq 0$, the empirical likelihood estimators reduce the asymptotic mean squared errors. In this section, we will relax this point identifying restriction, and allow θ_0 to be partially identified by model (4.1).

The identified region is defined by

$$\Omega = \{\theta : Em_1(X,\theta) \geq 0, Em_2(X,\theta) = 0\}$$

In our setting, Ω shrinks to a lower dimension sub-manifold of $\{\theta : Em_2(X,\theta) = 0\}$ with boundaries defined by linear or nonlinear hyperplanes $\{\theta : Em_1(X,\theta) = 0\}$. One of the problem one needs to take into account when considering the asymptotic behaviors of the posterior distribution is that Ω has zero Lebesgue measure, due to the loss of dimensionality. Thus integrating over Ω is always zero. The limit of posterior density is known as Dirac function:

$$\lim_{n\to\infty} p(\theta|X^n) = \begin{cases} +\infty & \theta \in \Omega \\ 0, & \theta \notin \Omega \end{cases} a.s.$$

Thus $\lim_{n\to\infty} p(\theta|X^n)$ is not a real valued function of θ.

However, it's still possible to study the large sample properties of the posterior distributions completely on Θ. Like in $int(\Omega) \neq \phi$ case, a dense subset in Ω plays an important role in characterizing such behaviors. Define

$$\Xi = \{\theta \in \Omega : Em_1(X,\theta) > 0\} \tag{1.2}$$

We will assume Ξ is dense in Ω and will study the large sample properties of the posterior around Ξ.

1.1 Derivation for Limited Information Likelihood

Suppose $X^n = \{X_1, ..., X_n\}$ is a stationary realization of X. Define $\bar{m}_j(\theta) = \frac{1}{n}\sum_{i=1}^n m_j(X_i, \theta)$, for $j = 1, 2$. Like before, we introduce auxiliary parameter λ to moment inequalities and define

$$G(\theta, \lambda) = \begin{pmatrix} \bar{m}_1(\theta) - \lambda \\ \bar{m}_2(\theta) \end{pmatrix}, \theta \in \Theta, \lambda \in [0, \infty)^r$$

For any positive definite $r \times r$ matrix V not depending on θ, define limited information likelihood:

$$L(\theta) = \int_{[0,\infty)^r} \frac{1}{\sqrt{\det(\frac{2\pi V}{n})}} e^{-\frac{n}{2}G(\theta,\lambda)^T V^{-1} G(\theta,\lambda)} p(\lambda) d\lambda$$

Write V^{-1} into subblocks

$$V^{-1} = \begin{pmatrix} \Sigma_1 & \Sigma_3 \\ \Sigma_3^T & \Sigma_2 \end{pmatrix}, \Sigma_1 : r \times r, \Sigma_2 : p \times p$$

We still place an exponential prior on λ:

$$p(\lambda) = (\prod_{i=1}^r \psi_i) e^{-\psi^T \lambda}, \psi, \lambda \in [0, \infty)^r$$

then we have

$$\begin{aligned} L(\theta) &= \int_{[0,\infty)^r} \frac{1}{\sqrt{\det(\frac{2\pi V}{n})}} \exp\left(-\frac{n}{2}(\bar{m}_1(\theta) - \lambda, \bar{m}_2(\theta)) \begin{pmatrix} \Sigma_1 & \Sigma_3 \\ \Sigma_3^T & \Sigma_2 \end{pmatrix} \begin{pmatrix} \bar{m}_1(\theta) - \lambda \\ \bar{m}_2(\theta) \end{pmatrix} \right) p(\lambda) d\lambda \\ &= \frac{\prod_{i=1}^r \psi_i}{\sqrt{\det(V_2)}} P(Z \geq 0) e^{\tau} \end{aligned}$$

where:

- Z follows multivariate normal distribution with mean μ, variance covariance matrix $\frac{\Sigma_1^{-1}}{n}$, $\mu = \bar{m}_1(\theta) + \Sigma_1^{-1}\Sigma_3^T \bar{m}_2(\theta) - \frac{1}{n}\Sigma_1^{-1}\psi$.
- $V_2 = (\Sigma_2 - \Sigma_3^T \Sigma_1^{-1} \Sigma_3)^{-1}$. If $V = Var(m_1, m_2)$, then by the matrix inversion formula, $V_2 = Var(m_2)$.
- $\tau = -\frac{n}{2}\bar{m}_2(\theta)^T V_2^{-1} \bar{m}_2(\theta) - \psi^T(\Sigma_1^{-1}\Sigma_3^T \bar{m}_2(\theta) + \bar{m}_1(\theta)) + \frac{1}{2n}\psi^T \Sigma_1^{-1} \psi$

Roughly speaking, when $\theta \notin \Omega$, either $Em_2(X, \theta) \neq 0$ or $\exists Em_{1j}(X, \theta) < 0$. When $Em_2(X, \theta) \neq 0$, since V_2^{-1} is also positive definite, $e^{\tau} \to 0$; when $Em_2(X, \theta) = 0$ but $Em_{1j}(X, \theta) < 0$ for some

j, then for large n, the jth component of $\mu < 0$. Since the covariance matrix of Z has order $O(n^{-1})$, $P(Z \geq 0) \to 0$. Therefore, $L(\theta) \to 0$ outside Ω.

When $\theta \in \Omega$, by central limit theorem, $\bar{m}_2(\theta) = O_p(n^{-1/2})$, hence $e^{\tau} = O_p(1)$. In addition, for large n, $P(Z \geq 0) \approx 1$. Thus $L(\theta) = O_p(1)$.

1.2 Posterior Distribution

Let $p(\theta)$ denote the prior on θ, then $p(\theta|X^n) \propto p(\theta)L(\theta)$. We will look at the posterior distribution, especially the convergence rate on the dense subset Ξ and $(\Omega^c)^{-\delta}$ for small enough δ.

Assumption 1.1. *Ξ defined in (1.2) is dense in Ω.*

This assumption states that if θ_0 satisfies $Em_2(X, \theta_0) = 0$ and $Em_{1j}(X, \theta_0) = 0$ for some $j = 1, ..., r$, then in any neighborhood of θ_0 we can find θ_1 such that $Em_1(X, \theta_1) > 0$ and $Em_2(X, \theta_1) = 0$.

Suppose all the other components of $Em_1(X, \theta_0)$ except for j are positive. By continuity of $Em_1(X, .)$, they remain to be positive in a small neighborhood of θ_0. Suppose Assumption 4.1 doesn't hold, then within some neighborhood U of θ_0, $Em_1(X, \theta) \leq 0$. Since Ω is connected, we argue that $Em_{1j}(X, \theta) \equiv 0$ on $U \cap \Omega$. Hence intuitively, Assumption 4.1 says that for each i, hyperplane $\{\theta : Em_{1i}(X, \theta) = 0\}$ has no part that overlaps with $\{\theta : Em_2(X, \theta) = 0\}$.

Example 1.1. This example shows Assumption 1.1 is satisfied by the interval regression model. Suppose we have moment inequalities $E(Z_1 Y_1) \leq E(Z_1 X^T)\theta \leq E(Z_1 Y_2)$ and exact moment condition $EZ_2(Y_3 - X^T\theta) = 0$, where Z_i, $i = 1, 2$ are r_1 and r_2 dimensional vectors of instrumental variables respectively, with each instrument being positive almost surely, and don't share same components. Y_i is scalar $i = 1, 2, 3$, and $\theta \in \mathbb{R}^d$. $Y_2 > Y_3 > Y_1$ a.s. Let $W = (Z_1, Z_2, X, Y_1, Y_2, Y_3)$, then

$$m_1(W, \theta) = \begin{pmatrix} Z_1(Y_2 - X^T\theta) \\ Z_1(X^T\theta - Y_1) \end{pmatrix}, m_2(W, \theta) = Z_2(Y_3 - X^T\theta)$$

We assume $r_2 < d$ so that θ can not be point identified by $Em_2(W, \theta) = 0$. Let's also assume $\exists$ a unit vector δ such that $EZ_2 X^T \delta = 0$ but $EZ_{11} X^T \delta < 0$, where Z_{11} denotes the first component of Z_1. In this interval instrumental variable regression model,

$$\Xi = \{\theta : E(Z_1 Y_1) < E(Z_1 X^T)\theta < E(Z_1 Y_2); EZ_2 Y_3 = EZ_2 X^T \theta\}$$

We now show Ξ is dense.

Pick up $\theta_0 \in \Omega \backslash \Xi$ such that $EZ_2(Y_3 - X^T\theta_0) = 0$, $EZ_{11}(Y_2 - X^T\theta_0) = 0$. Let's assume $Em_{1j}(W, \theta_0) > 0$ for $j > 1$ for simplicity. Then in a small neighborhood of θ_0, $Em_{1j}(W, .) > 0$ for $j > 1$. For small enough $\epsilon > 0$, let $\theta_1 = \theta_0 + \epsilon\delta$, then

$$Em_2(W, \theta_1) = EZ_2(Y_3 - X^T\theta_0) - \epsilon EZ_2 X^T \delta = 0$$

$$Em_{11}(W, \theta_1) = EZ_{11}(Y_2 - X^T\theta_0) - \epsilon EZ_{11} X^T \delta = -\epsilon EZ_{11} X^T \delta > 0$$

Therefore $\theta_1 \in B(\theta_0, 2\epsilon) \cap \Xi$.

Assumption 1.2. *(i) $Em_{1j}(X, \theta)$ is continuous on Θ for each j.*

(ii) $Em_{2j}(X, \theta)$ is Lipschitz continuous on Θ for each j.

Assumption 1.3. *w.p.a.1, for any $\beta_n \to \infty$,*

$$\sup_{\theta \in \Theta} ||\bar{m}_2(\theta) - Em_2(X, \theta)||^2 \leq \frac{\ln \beta_n}{n}$$

Assumption 1.4. *$p(\theta)$ is continuous, and bounded away from zero and infinity on Ω.*

Theorem 1.1. *Under Assumption 2.1, 2.2 in Liao and Jiang (2008), and 1.1-1.4, then*

1. *$\forall \delta > 0$, for some $\alpha > 0$,*

$$P(\theta \in (\Omega^c)^{-\delta} | X^n) = o_p(e^{-\alpha n})$$

2. *$\forall \omega \in \Xi$, $\exists R > 0$, $\forall \delta < R$, for all $\beta_n \to \infty$, we have in probability*

$$P(\theta \in B(\omega, \delta) | X^n) \succ \frac{1}{\beta_n} n^{-d/2}$$

where $d = \dim(\omega)$

Like the case when $int(\Omega) \neq \phi$, let $g(.)$ be a continuous real-valued function on Θ, let $F_g^{-1}(y)$ be the $y-$quantile of the posterior cdf of $g(\theta)$.

Theorem 1.2. *Under Assumption 2.1-2.3 in Liao and Jiang (2008) and 1.1-1.3, if* $\{\pi_n\}_{n=1}^{\infty}$ *is such that* $e^{-\alpha n} \prec \pi_n \prec n^{-\beta}$*, for any* $\alpha > 0$ *and some* $\beta > \frac{d}{2}$*, then*

$$d_H([F_g^{-1}(\pi_n), F_g^{-1}(1-\pi_n)], g(\Omega)) \to 0 \text{ in probability.}$$

2 Monte Carlo Experiments

We simulate an interval instrumental regression model with exact moment conditions.

Example 2.1 (Interval regression models with exact moment condition)**.** We consider,

$$E(Z_1Y_1) \le E(Z_1X^T)\theta \le E(Z_1Y_2), \qquad E(Z_2X^T)\theta = E(Z_2Y_3)$$

We generate $(X_1, X_2) \sim N_2((1,1)^T, I_2)$, and $Z_1 = X_1 + X_2$, $Z_2 = -2X_1 + 2X_2$. $(Y_1, Y_2)^T \sim N_2((3,6)^T, 0.1I_2)$, independent of X. Let $Y_3 = Z_1 + 3$. Then the identified region is given by

$$\Omega = \{(\theta_1, \theta_2) : \theta_1 = \theta_2, 2\theta_1 + \theta_2 \le 4\}$$

To estimate Ω, we choose a positive definite weight matrix

$$V^{-1} = \begin{pmatrix} I_2 & \Sigma_3 \\ \Sigma_3^T & 6 \end{pmatrix}$$

where $\Sigma_3 = (1,2)^T$.

Figure 1 displays the identified region as well as 10,000 draws using Metropolis algorithm, with two choices of $\psi^1 = (0.5, 0.5)^T$, and $\psi^2 = (0.01, 0.01)^T$ respectively.

We also estimate the identified interval of θ_1, which is [1, 2] theoretically. Table 3 reports $[F_e^{-1}(\pi_n), F_e^{-1}(1-\pi_n)]$ based on the empirical cdf F_e of 5000 draws from the posterior distribution.

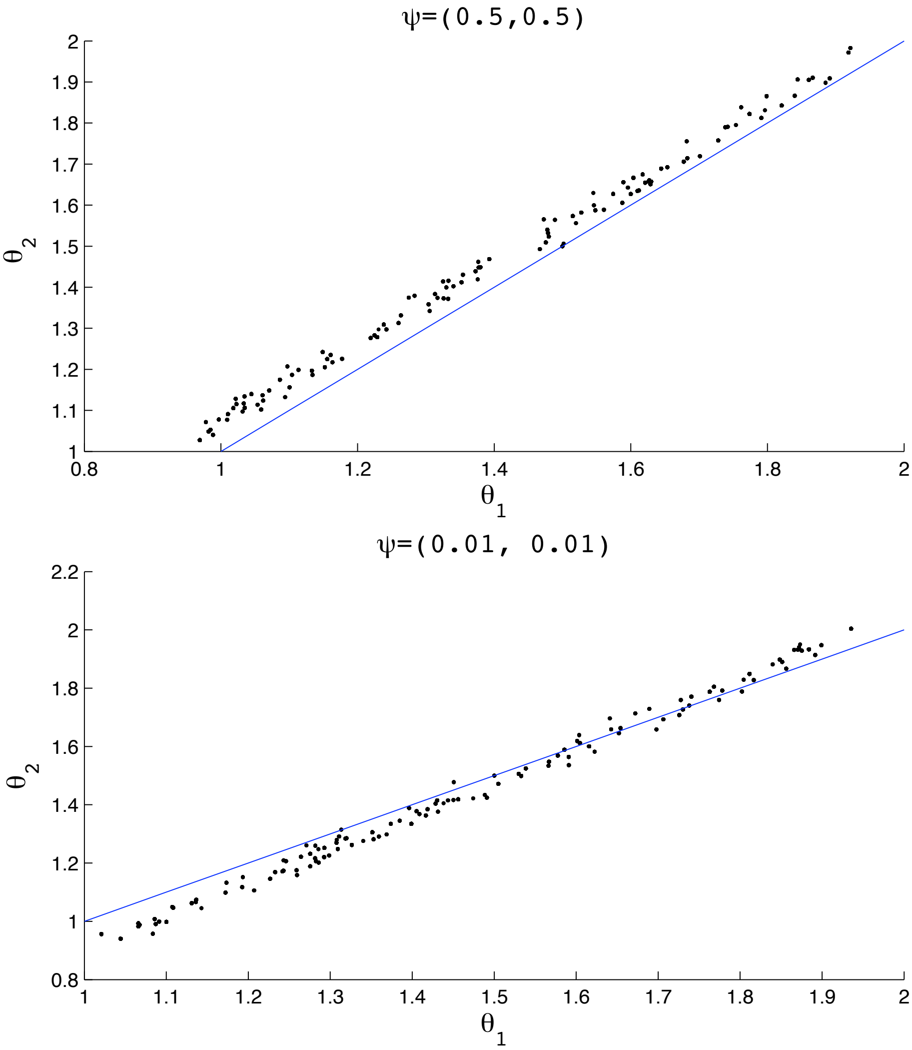

Figure 1: The identified set and MCMC draws

Table 1: Estimation of $\Omega_1 = [1, 2]$ based on the empirical cdf

π_n	$e^{-\sqrt{n}}$	$\frac{1}{\sqrt{n}}$	$\frac{1}{\ln n}$
$n = 500$	[1.1384, 2.0295]	[1.0068, 1.9331]	[1.0904, 1.6207]
$n = 1000$	[1.0809, 1.9425]	[0.9620, 1.8844]	[1.1183, 1.8874]
$n = 5000$	[1.1045, 1.8551]	[0.9944, 1.9575]	[1.1878, 1.9729]

3 Proofs

Define $A_\delta = \{\theta : Em_2(X,\theta)^T V_2^{-1} Em_2(X,\theta) > \delta\}$, and

$$A_2 = \left\{\theta : Em_2(X,\theta) = 0, \min_j Em_{1j}(X,\theta) < 0\right\}$$

Lemma 3.1. $\forall \delta > 0$, *for some* $a > 0$

$$\int_{A_\delta \cup A_2} p(\theta)L(\theta)d\theta = o_p(e^{-an})$$

Proof. Define $\hat{A}_\delta = \{\theta : \bar{m}_2(\theta)^T V_2^{-1} \bar{m}_2(\theta) > \delta\}$,

$$\begin{aligned}\int_{A_\delta} p(\theta)L(\theta)d\theta &= \int_{A_\delta \cap \hat{A}_\delta} p(\theta)L(\theta)d\theta + \int_{A_\delta \cap \hat{A}_\delta^c} p(\theta)L(\theta)d\theta \\ &\leq \int_{\hat{A}_\delta} p(\theta)L(\theta)d\theta + \int_{A_\delta \cap \hat{A}_\delta^c} p(\theta)L(\theta)d\theta\end{aligned}$$

$A_\delta \cap \hat{A}_\delta^c = \{\theta : Em_2(X,\theta)^T V_2^{-1} Em_2(X,\theta) > \delta\} \cap \{\theta : \bar{m}_2(\theta)^T V_2^{-1} \bar{m}_2(\theta) \leq \delta\} \to \phi$ w.p.a.1. Hence for large n, $\mu(A_\delta \cap \hat{A}_\delta^c) = 0$. Then $\exists N$, when $n > N$, w.p.a.1,

$$\int_{A_\delta} p(\theta)L(\theta)d\theta \leq \int_{\hat{A}_\delta} p(\theta)\frac{\prod_i \psi_i}{\sqrt{\det(V_2)}} e^{-\frac{n}{2}\bar{m}_2(\theta)^T V_2^{-1} \bar{m}_2(\theta) - \psi^T(\Sigma_1^{-1}\Sigma_3^T \bar{m}_2(\theta) + \bar{m}_1(\theta)) + \frac{1}{2n}\psi^T \Sigma_1^{-1}\psi} d\theta$$

For some $\epsilon > 0$, for large n,

$$e^{-\psi^T(\Sigma_1^{-1}\Sigma_3^T \bar{m}_2(\theta) + \bar{m}_1(\theta)) + \frac{1}{2n}\psi^T\Sigma_1^{-1}\psi} \leq e^{\|\psi\|(\sup_{\theta\in\Theta}\|\Sigma_1^{-1}\Sigma_3^T Em_2(X,\theta) + Em_1(X,\theta)\| + \epsilon) + \epsilon} < \infty$$

Thus for some positive constant C, and large n,

$$\int_{A_\delta} p(\theta)L(\theta)d\theta \leq C \cdot \int_{\hat{A}_\delta} p(\theta) e^{-\frac{n}{2}\bar{m}_2(\theta)^T V_2^{-1}\bar{m}_2(\theta)} d\theta \leq C \cdot e^{-\frac{\delta}{2}n}$$

In addition, $\mu(A_2) = 0$ and $p(\theta)L(\theta)$ is bounded on Θ, hence

$$\int_{A_\delta \cup A_2} p(\theta)L(\theta)d\theta \leq \int_{A_\delta} p(\theta)L(\theta)d\theta + \int_{A_2} p(\theta)L(\theta)d\theta = O_p(e^{-\frac{\delta}{2}n})$$

□

Lemma 3.2. $\forall \delta > 0$, *for some* $a > 0$,

$$\int_{(\Omega^c)^{-\delta}} p(\theta)L(\theta)d\theta = o_p(e^{-an})$$

Proof. : $\forall\theta \in (\Omega^c)^{-\delta}$, then either $\exists\delta(\theta) > 0, \theta \in A_{\delta(\theta)}$, or $\theta \in A_2$, hence $\theta \in A_{\delta(\theta)} \cup A_2$. Thus $(\Omega^c)^{-\delta} \subset \bigcup_{\theta\in(\Omega^c)^{-\delta}}[A_{\delta(\theta)} \cup A_2]$. Note that $(\Omega^c)^{-\delta} = \{\theta : d(\theta, \Omega) \geq \delta\}$ is compact, hence $\exists$ $\{A_{\delta 1} \cup A_2, ..., A_{\delta N} \cup A_2\} \subset \{A_{\delta(\theta)} \cup A_2 : \theta \in (\Omega^c)^{-\delta}\}$ such that

$$(\Omega^c)^{-\delta} \subset \bigcup_{i=1}^{N}[A_{\delta i} \cup A_2]$$

Let $\delta^* = \min\{\delta_i, i = 1, ..., N\}$. For $a > b > 0$, $A_a \subset A_b$, hence $(\Omega^c)^{-\delta} \subset A_{\delta^*} \cup A_2$. Therefore

$$\int_{(\Omega^c)^{-\delta}} p(\theta)L(\theta)d\theta \leq \int_{A_{\delta^*}\cup A_2} p(\theta)L(\theta)d\theta = o_p(a^{-an})$$

□

Lemma 3.3. *If Z_θ follows $N_r(\bar{m}_1(\theta) + \Sigma_1^{-1}\Sigma_3^T\bar{m}_2(\theta) - \frac{1}{n}\Sigma_1^{-1}\psi, \frac{1}{n}\Sigma_1^{-1})$, then $\forall\omega \in \Xi$, $\exists R > 0$, w.p.a.1,*

$$\liminf_{n\to\infty} \inf_{\theta\in B(\omega,R)} P(Z_\theta \geq 0) > 0$$

Proof. Let $\xi_n(\theta) = \bar{m}_1(\theta) + \Sigma_1^{-1}\Sigma_3^T\bar{m}_2(\theta) - \frac{1}{n}\Sigma_1^{-1}\psi$. $\forall\omega \in \Xi$, $Em_1(X, \omega) > 0$. Since $Em_1(X, \theta)$ is continuous on Θ, there exist $\epsilon > 0$, and an open ball $B(\omega, R_1)$, such that $\inf_{\theta\in B(\omega,R_1)} Em_1(X, \theta) > \epsilon$, where the inequality is taken coordinately. Moreover, $Em_2(X, \omega) = 0$; hence by the continuity of $Em_2(X, .)$, $\exists R < R_1$ such that $\sup_{\theta\in B(\omega,R)} |\Sigma_1^{-1}\Sigma_3^T Em_2(X, \theta)| < \epsilon$, where $|.|$ denotes the absolute value, taken coordinately. Therefore, $\inf_{\theta\in B(\omega,R)}(Em_1(X, \theta) + \Sigma_1^{-1}\Sigma_3^T Em_2(X, \theta)) > 0$. $\exists N$, when $n > N$, w.p.a.1, coordinately.

$$\inf_{\theta\in B(\omega,R)}(\bar{m}_1(\theta) + \Sigma_1^{-1}\Sigma_3^T\bar{m}_2(\theta) - \frac{1}{n}\Sigma_1^{-1}\psi) = \inf_{\theta\in B(\omega,R)} \xi_n(\theta) > 0$$

Let σ_{1j}^2 denote the jth diagonal element in Σ_1^{-1}, and $\xi_{nj}(\theta)$ denote the jth element of $\xi_n(\theta)$. Then

$$\begin{aligned}\inf_{\theta\in B(\omega,R)} P(Z_\theta \geq 0) \quad &\geq \quad 1 - r \cdot \Phi\left(-\sqrt{n} \inf_{\theta\in B(\omega,R)} \min_j \frac{\xi_{nj}(\theta)}{\sqrt{\sigma_{1j}^2}}\right) \\ &\geq \quad 1 - r \cdot \Phi\left(-\sqrt{n} \min_j \frac{\inf_{\theta\in B(\omega,R)} \xi_{nj}(\theta)}{\sqrt{\sigma_{1j}^2}}\right) \\ &>_n \quad 0\end{aligned}$$

□

Lemma 3.4. *For any $\beta_n \to \infty$, $\forall\omega \in \Xi$, $\exists R > 0$, $\forall\delta < R$, w.p.a.1,*

$$\int_{B(\omega,\delta)} p(\theta)L(\theta)d\theta \succ \frac{1}{\beta_n}n^{-d/2}$$

where $d = \dim(\omega)$.

Proof. $\forall \omega \in \Xi$, it can be shown that (using lemma C.3), $\exists R > 0$, and a positive constant C, such that

$$\int_{B(\omega,R)} p(\theta)L(\theta)d\theta \geq C \int_{B(\omega,R)} p(\theta)e^{-\frac{n}{2}\bar{m}_2(\theta)^T V_2^{-1}\bar{m}_2(\theta)}d\theta$$

For deterministic V_2^{-1}, and a vector α, we write weighted norm $\|\alpha\|_V^2 = \alpha^T V_2^{-1}\alpha$. Then we have

$$\frac{1}{2}\|\bar{m}_2(\theta)\|_V^2 \leq \|Em_2(X,\theta)\|_V^2 + \|\bar{m}_2(\theta) - Em_2(X,\theta)\|_V^2$$

By assumption 4.3, for any $\beta_n \to \infty$, choose $\beta_n^{\|V^{-1}\|^{-1}}$, so that

$$e^{-n\|\bar{m}_2(\theta)-Em_2(X,\theta)\|_V^2} \geq e^{-\ln \beta_n^{\|V^{-1}\|^{-1}}\|V^{-1}\|} = \beta_n^{-\|V^{-1}\|^{-1}\cdot\|V^{-1}\|} = \frac{1}{\beta_n}$$

For some constant $\alpha > 1$, let $U = \{\theta : \|Em_2(X,\theta)\|_V^2 < \frac{\ln \alpha}{n}\}$. By assumption 3.3, $\exists R' < R$, such that for any $0 < \delta < R'$, $\inf_{\theta \in B(\omega,\delta)} p(\theta) > 0$. Then

$$\int_{B(\omega,\delta)} p(\theta)L(\theta)d\theta \geq C \int_{B(\omega,\delta)\cap U} \frac{1}{\alpha\beta_n}p(\theta)d\theta \geq \frac{Const}{\beta_n}\mu(B(\omega,\delta)\cap U)$$

To derive a lower bound for the Lebesgue measure of $B(\omega,\delta)\cap U$, note that $Em_2(X,\theta)$ is Lipschitz continuous, and $Em_2(X,\omega) = 0$, $\exists\ \lambda > 0$, such that $\forall \theta \in B(\omega,\delta)$, $\|Em_2(X,\theta)\|^2 \leq \lambda\|\theta-\omega\|^2$. Then

$$\begin{aligned}\|Em_2(X,\theta)\|_V^2 &\leq \|Em_2(X,\theta)\|^2 \cdot \|V^{-1}\|^2 \\ &\leq \lambda\|V^{-1}\|^2 \cdot \|\theta-\omega\|^2 \\ &\leq \lambda\|V^{-1}\|^2|\theta-\omega|_\infty^2\end{aligned}$$

where $|\theta-\omega|_\infty = \max_j |\theta_j - \omega_j|$. Hence $\{\theta : |\theta-\omega|_\infty^2 < \frac{\ln \alpha}{\lambda\|V^{-1}\|^2 n}\} \subset U$. Moreover, for large enough n, $\{\theta : |\theta-\omega|_\infty^2 < \frac{\ln \alpha}{\lambda\|V^{-1}\|^2 n}\} \subset B(\omega,\delta)$, thus $\mu(B(\omega,\delta)\cap U) \geq \mu(\{\theta : |\theta-\omega|_\infty^2 < \frac{\ln \alpha}{\lambda\|V^{-1}\|^2 n}\}) = (2\sqrt{\frac{\ln \alpha}{\lambda\|V^{-1}\|^2}})^d n^{-d/2}$. Hence $\int_{B(\omega,\delta)} p(\theta)L(\theta)d\theta \succ \frac{1}{\beta_n}n^{-d/2}$.

Proof of Theorem 1.1

Proof.

1. Let $\beta_n = n^{d/2}$. Lemma C.4 implies that $\int_\Theta p(\theta)L(\theta)d\theta \succ n^{-d}$. Thus by Lemma C.2, for some $\alpha > 0$,

$$P(\theta \in (\Omega^c)^{-\delta}|X^n) = \frac{\int_{(\Omega^c)^{-\delta}} p(\theta)L(\theta)d\theta}{\int_\Theta p(\theta)L(\theta)d\theta} \prec \frac{o_p(e^{-\alpha n})}{n^{-d}} = o_p(e^{-\frac{\alpha}{2}n})$$

2. The result follows immediately from Lemma C.4 and that $\int_\Theta p(\theta)L(\theta)d\theta$ is bounded. □

References

[1] Liao, Y. and Jiang, W. (2008). Bayesian analysis in moment inequality models. *working paper*. Northwestern University.

[2] Moon, H. and Schorfheide, F. (2006), Boosting Your Instruments: Estimation with Overidentifying Inequality Moment Conditions. CEPR Discussion Paper No. 5605 Simple Measure, *Review of Economics and Statistics*. LXXIX, 348-352.