A uniform Berry-Esseen theorem on $M$–estimators for geometrically ergodic Markov chains

Valentin Patilea

IRMAR-INSa & CREST-ENSAI, Rennes, France

en collaboration avec
Loïc Hervé & James Ledoux, IRMAR-INSa, Rennes

Colloque Franco-Roumain, Poitiers, Août 2010
Let \( \{X_n\}_{n \geq 0} \) be a \( V \)-geometrically ergodic Markov chain with \( V(\cdot) \geq 1 \) some fixed unbounded real-valued function.
The problem (1/2)

- Let \( \{X_n\}_{n \geq 0} \) be a \( V \)-geometrically ergodic Markov chain with \( V(\cdot) \geq 1 \) some fixed unbounded real-valued function.

- Consider

\[
M_n(\alpha) = n^{-1} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k),
\]

\( \alpha \in \mathcal{A} \subset \mathbb{R} \) with \( F(\cdot, \cdot, \cdot) \) real-valued functional and \( \mathcal{A} \) some open interval.
Let \( \{X_n\}_{n \geq 0} \) be a \( V \)-geometrically ergodic Markov chain with \( V(\cdot) \geq 1 \) some fixed unbounded real-valued function.

Consider

\[
M_n(\alpha) = n^{-1} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k),
\]

\( \alpha \in A \subset \mathbb{R} \) with \( F(\cdot, \cdot, \cdot) \) real-valued functional and \( A \) some open interval.

Define the \( M \)-estimator \( \hat{\alpha}_n \) such that

\[
M_n(\hat{\alpha}_n) \leq \inf_{\alpha \in A} M_n(\alpha) + c_n
\]

with \( c_n, n \geq 1 \) some sequence of real numbers decreasing to zero.
Let \( \{X_n\}_{n \geq 0} \) be a \( V \)-geometrically ergodic Markov chain with \( V(\cdot) \geq 1 \) some fixed unbounded real-valued function.

Consider

\[
M_n(\alpha) = n^{-1} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k),
\]

\( \alpha \in A \subset \mathbb{R} \) with \( F(\cdot, \cdot, \cdot) \) real-valued functional and \( A \) some open interval.

Define the \( M \)-estimator \( \hat{\alpha}_n \) such that

\[
M_n(\hat{\alpha}_n) \leq \inf_{\alpha \in A} M_n(\alpha) + c_n
\]

with \( c_n, \ n \geq 1 \) some sequence of real numbers decreasing to zero.

Let \( \alpha_0 \in A \) be the “true” value of the parameter of interest.
Suppose that

$$\left( \left| \frac{\partial F}{\partial \alpha}(\alpha, x, y) \right| + \left| \frac{\partial^2 F}{\partial \alpha^2}(\alpha, x, y) \right| \right)^{3+\varepsilon} \leq C \left( V(x) + V(y) \right)$$

for some constants $\varepsilon > 0$ and $C > 0$. 
Suppose that
\[
\left( \left| \frac{\partial F}{\partial \alpha}(\alpha, x, y) \right| + \left| \frac{\partial^2 F}{\partial \alpha^2}(\alpha, x, y) \right| \right)^{3+\varepsilon} \leq C \left( V(x) + V(y) \right)
\]
for some constants $\varepsilon > 0$ and $C > 0$.

**Aim:** show that the estimator $\hat{\alpha}_n$ satisfies a Berry-Esseen theorem uniformly with respect to the underlying probability distribution of the Markov chain.
Consider the observations are generated by the process

\[ X_n = \rho_0 X_{n-1} + \sigma(X_{n-1}; \zeta_0, \gamma_0) \varepsilon_n, \quad n = 1, 2, \ldots \]
Example: AR(1) with ARCH(1) errors

Consider the observations are generated by the process

\[ X_n = \rho_0 X_{n-1} + \sigma(X_{n-1}; \zeta_0, \gamma_0) \epsilon_n, \quad n = 1, 2, \ldots \]

- \( X_0 = 0, \sigma^2(x; \zeta, \gamma) = \zeta + \gamma x^2; \)
- \( \{\epsilon_n\}_{n \geq 1} \) are i.i.d. zero mean of variance equal to 1, with finite \( p \)th order moment, \( p > 0 \) and (unknown) continuous density \( f_\epsilon > 0; \)
- \(|\rho_0| < 1\) and \( \zeta_0, \gamma_0 > 0 \) are the true values of the parameters.
Example: AR(1) with ARCH(1) errors

Consider the observations are generated by the process

\[ X_n = \rho_0 X_{n-1} + \sigma(X_{n-1}; \zeta_0, \gamma_0) \varepsilon_n, \quad n = 1, 2, \ldots \]

- \( X_0 = 0, \sigma^2(x; \zeta, \gamma) = \zeta + \gamma x^2; \)
- \( \{\varepsilon_n\}_{n \geq 1} \) are i.i.d. zero mean of variance equal to 1, with finite \( p \)th order moment, \( p > 0 \) and (unknown) continuous density \( f_\varepsilon > 0; \)
- \( |\rho_0| < 1 \) and \( \zeta_0, \gamma_0 > 0 \) are the true values of the parameters.

The “true” parameter is

\[ \theta = (\rho_0, \zeta_0, \gamma_0) \in \Theta \subset [-\bar{\rho}, \bar{\rho}] \times [m_\zeta, M_\zeta] \times [m_\gamma, M_\gamma] \subset \mathbb{R}^3, \text{ where} \]
Example: AR(1) with ARCH(1) errors

Consider the observations are generated by the process

\[ X_n = \rho_0 X_{n-1} + \sigma(X_{n-1}; \zeta_0, \gamma_0) \varepsilon_n, \quad n = 1, 2, \ldots \]

- \(X_0 = 0, \sigma^2(x; \zeta, \gamma) = \zeta + \gamma x^2;\)
- \(\{\varepsilon_n\}_{n \geq 1}\) are i.i.d. zero mean of variance equal to 1, with finite \(p\)th order moment, \(p > 0\) and (unknown) continuous density \(f_\varepsilon > 0;\)
- \(|\rho_0| < 1\) and \(\zeta_0, \gamma_0 > 0\) are the true values of the parameters.

The “true” parameter is

\[ \theta = (\rho_0, \zeta_0, \gamma_0) \in \Theta \subset [-\bar{\rho}, \bar{\rho}] \times [m_\zeta, M_\zeta] \times [m_\gamma, M_\gamma] \subset \mathbb{R}^3, \]

- \(\bar{\rho} \in (0, 1)\) is given;
- \(0 < m_\zeta < M_\zeta < \infty\) and \(0 < m_\gamma < M_\gamma < 1\) are given such that \(\bar{\rho} + \sqrt{M_\gamma} < 1.\)
Example: AR(1) with ARCH(1) errors

Consider the observations are generated by the process

\[ X_n = \rho_0 X_{n-1} + \sigma(X_{n-1}; \zeta_0, \gamma_0) \varepsilon_n, \quad n = 1, 2, \ldots \]

- \( X_0 = 0, \sigma^2(x; \zeta, \gamma) = \zeta + \gamma x^2; \)
- \( \{\varepsilon_n\}_{n \geq 1} \) are i.i.d. zero mean of variance equal to 1, with finite \( p \)th order moment, \( p > 0 \) and (unknown) continuous density \( f_{\varepsilon} > 0; \)
- \( |\rho_0| < 1 \) and \( \zeta_0, \gamma_0 > 0 \) are the true values of the parameters.

The “true” parameter is

\[ \theta = (\rho_0, \zeta_0, \gamma_0) \in \Theta \subset [\bar{\rho}, \rho] \times [m_{\zeta}, M_{\zeta}] \times [m_{\gamma}, M_{\gamma}] \subset \mathbb{R}^3, \] where
- \( \bar{\rho} \in (0, 1) \) is given;
- \( 0 < m_{\zeta} < M_{\zeta} < \infty \) and \( 0 < m_{\gamma} < M_{\gamma} < 1 \) are given such that \( \bar{\rho} + \sqrt{M_{\gamma}} < 1. \)

We are interested to estimate \( \rho_0 \) and \( \gamma_0 \) and to derive BE bounds.
To estimate \( \rho_0 \) one can use the least squares estimator

\[
\hat{\rho}_n = \frac{\sum_{k=1}^{n} X_k X_{k-1}}{\sum_{k=1}^{n} X_{k-1}^2} = \arg \min_{\rho} \frac{1}{n} \sum_{k=1}^{n} F(\rho, X_{k-1}, X_k),
\]

where \( F(\rho, X_{k-1}, X_k) = (X_k - \rho X_{k-1})^2 \).
To estimate $\rho_0$ one can use the least squares estimator

$$\hat{\rho}_n = \frac{\sum_{k=1}^{n} X_k X_{k-1}}{\sum_{k=1}^{n} X_{k-1}^2} = \arg \min_{\rho} \frac{1}{n} \sum_{k=1}^{n} F(\rho, X_{k-1}, X_k),$$

where $F(\rho, X_{k-1}, X_k) = (X_k - \rho X_{k-1})^2$.

The parameter $\gamma_0$ can also be estimated by a suitable least squares approach.
Agenda of the talk

1. Introduction
Agenda of the talk

1. Introduction
2. A general method for deriving Berry-Esseen bounds for $M$-estimators
3. A uniform Berry-Esseen statement with geometrically ergodic Markov chains
4. Application: Berry-Esseen bound for $M$-estimators with geometrically ergodic Markov chains
5. Example of application: AR(1) with ARCH(1) errors
Agenda of the talk

1. Introduction

2. A general method for deriving Berry-Esseen bounds for $M$-estimators

3. A uniform Berry-Esseen statement with geometrically ergodic Markov chains
Agenda of the talk

1. Introduction
2. A general method for deriving Berry-Esseen bounds for $M$—estimators
3. A uniform Berry-Esseen statement with geometrically ergodic Markov chains
4. Application: Berry-Esseen bound for $M$—estimators with geometrically Markov chains
1. Introduction

2. A general method for deriving Berry-Esseen bounds for $M$–estimators

3. A uniform Berry-Esseen statement with geometrically ergodic Markov chains


5. Example of application: AR(1) with ARCH(1) errors
Introduction

A general method for deriving Berry-Esseen bounds for $M$–estimators

A uniform Berry-Esseen statement with geometrically ergodic Markov chains

Application: Berry-Esseen bound for $M$–estimators with geometrically Markov chains

Example of application: AR(1) with ARCH(1) errors
Let $E$ be a measurable space equipped with a countably generated $\sigma$-field $\mathcal{E}$.
Let $E$ be a measurable space equipped with a countably generated $\sigma$-field $\mathcal{E}$.

Let $\{X_n\}_{n \geq 0}$ be a Markov chain with state space $E$ and transition kernels $\{Q_\theta(x, \cdot) : x \in E\}$ where $\theta$ is a parameter in some general set $\Theta$. 

Let $E$ be a measurable space equipped with a countably generated $\sigma$-field $\mathcal{E}$.

Let $\{X_n\}_{n \geq 0}$ be a Markov chain with state space $E$ and transition kernels $\{Q_\theta(x, \cdot) : x \in E\}$ where $\theta$ is a parameter in some general set $\Theta$.

Let $\mu$ be the initial distribution of the chain. It can be fixed or dependent on $\theta$. 
Let $E$ be a measurable space equipped with a countably generated $\sigma$-field $\mathcal{E}$.

Let $\{X_n\}_{n \geq 0}$ be a Markov chain with state space $E$ and transition kernels $\{Q_\theta(x, \cdot) : x \in E\}$ where $\theta$ is a parameter in some general set $\Theta$.

Let $\mu$ be the initial distribution of the chain. It can be fixed or dependent on $\theta$.

Let $\mathbb{P}_{\theta, \mu}$ be the probability distribution of $\{X_n\}_{n \geq 0}$ and $\mathbb{E}_{\theta, \mu}$ be the expectation w.r.t. $\mathbb{P}_{\theta, \mu}$.
For all \( \theta \), let

\[ M_\theta(\alpha) = \lim_{n \to \infty} \mathbb{E}_{\theta, \mu}[M_n(\alpha)]. \]

Assume that there exists a unique “true” value \( \alpha_0 = \alpha_0(\theta) \) of the parameter of interest, that is

\[ M_\theta(\alpha_0) < M_\theta(\alpha), \quad \forall \alpha \neq \alpha_0. \]
For all $\theta$, let

$$M_\theta(\alpha) = \lim_{n \to \infty} \mathbb{E}_{\theta, \mu}[M_n(\alpha)].$$

Assume that there exists a unique “true” value $\alpha_0 = \alpha_0(\theta)$ of the parameter of interest, that is

$$M_\theta(\alpha_0) < M_\theta(\alpha), \quad \forall \alpha \neq \alpha_0.$$

We want to prove

$$\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \mu} \left\{ \frac{\sqrt{n}}{\tau(\theta)} (\hat{\alpha}_n - \alpha_0) \leq u \right\} - \Gamma(u) \right| = O(n^{-1/2})$$

where
For all $\theta$, let

$$M_\theta(\alpha) = \lim_{n \to \infty} \mathbb{E}_{\theta, \mu}[M_n(\alpha)].$$

Assume that there exists a unique “true” value $\alpha_0 = \alpha_0(\theta)$ of the parameter of interest, that is

$$M_\theta(\alpha_0) < M_\theta(\alpha), \quad \forall \alpha \neq \alpha_0.$$

We want to prove

$$\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \mu} \left\{ \frac{\sqrt{n}}{\tau(\theta)} (\hat{\alpha}_n - \alpha_0) \leq u \right\} - \Gamma(u) \right| = O(n^{-1/2})$$

where

- $\Gamma(\cdot)$ denotes the $N(0, 1)$ d.f.
- $\tau(\theta)$ is some suitable positive real number.
Related work

- IID framework and functional $F(\alpha, X_k)$:
  - Pfanzagl (1971): $\{X_n\}_{n \geq 0}$ a sequence of i.i.d. random variables and with functionals of the form $F(\alpha, X_k)$.
  - Bentkus, Bloznelis & Götze (1997): obtain the result of Pfanzagl (with constants) using convexity arguments that allow for non-differentiable functionals $\alpha \mapsto F(\alpha, X_k)$.
Related work

- **IID framework and functional** $F(\alpha, X_k)$:
  - Pfanzagl (1971): $\{X_n\}_{n \geq 0}$ a sequence of i.i.d. random variables and with functionals of the form $F(\alpha, X_k)$.
  - Bentkus, Bloznelis & Götze (1997): obtain the result of Pfanzagl (with constants) using convexity arguments that allow for non differentiable functionals $\alpha \mapsto F(\alpha, X_k)$

- **Markov framework and functional** $F(\alpha, X_k, X_{k-1})$:
  - Rao (1973): extends the framework of Pfanzagl to $\{X_n\}_{n \geq 0}$ a uniformly ergodic Markov chain
  - Milhaud & Raugi (1989): ML in linear autoregressive models under strong conditions on the error term law. Moreover, they obtain a suboptimal rate $n^{-1/2} \ln^{1/2} n$. 
1 Introduction

2 A general method for deriving Berry-Esseen bounds for $M$–estimators

3 A uniform Berry-Esseen statement with geometrically ergodic Markov chains

4 Application: Berry-Esseen bound for $M$–estimators with geometrically Markov chains

5 Example of application: AR(1) with ARCH(1) errors
Pfanzagl’s method revisited

Consider a statistical model \((\Omega, \mathcal{F}, \{P_\theta, \theta \in \Theta\})\), where \(\Theta\) denotes some parameter space.
Consider a statistical model \((\Omega, \mathcal{F}, \{\mathbb{P}_\theta, \theta \in \Theta\})\), where \(\Theta\) denotes some parameter space.

Let \(\{X_n\}_{n \geq 0}\) be any sequence of observations (not necessarily markovian). Let \(\mathbb{E}_\theta\) denote the expectation with respect to \(\mathbb{P}_\theta\).
Consider a statistical model \((\Omega, \mathcal{F}, \{P_\theta, \theta \in \Theta\})\), where \(\Theta\) denotes some parameter space.

Let \(\{X_n\}_{n \geq 0}\) be any sequence of observations (not necessarily markovian). Let \(E_\theta\) denote the expectation with respect to \(P_\theta\).

For each \(n\), let \(M_n(\alpha)\) be a measurable function depending on \(X_0, X_1, \ldots, X_n\) and the parameter of interest \(\alpha \in A \subset \mathbb{R}\) with \(A\) some open interval.
Consider a statistical model \((\Omega, \mathcal{F}, \{\mathbb{P}_\theta, \theta \in \Theta\})\), where \(\Theta\) denotes some parameter space.

Let \(\{X_n\}_{n \geq 0}\) be any sequence of observations (not necessarily markovian). Let \(\mathbb{E}_\theta\) denote the expectation with respect to \(\mathbb{P}_\theta\).

For each \(n\), let \(M_n(\alpha)\) be a measurable function depending on \(X_0, X_1, ..., X_n\) and the parameter of interest \(\alpha \in A \subset \mathbb{R}\) with \(A\) some open interval.

Let \(c_n \downarrow 0\).
Pfanzagl’s method revisited

- Consider a statistical model \((\Omega, \mathcal{F}, \{P_\theta, \theta \in \Theta\})\), where \(\Theta\) denotes some parameter space.
- Let \(\{X_n\}_{n \geq 0}\) be any sequence of observations (not necessarily markovian). Let \(E_\theta\) denote the expectation with respect to \(P_\theta\).
- For each \(n\), let \(M_n(\alpha)\) be a measurable function depending on \(X_0, X_1, \ldots, X_n\) and the parameter of interest \(\alpha \in \mathcal{A} \subset \mathbb{R}\) with \(\mathcal{A}\) some open interval.
- Let \(c_n \downarrow 0\).
- **\(M-\text{estimator}\):** a measurable function \(\hat{\alpha}_n\) of the observations \((X_0, \ldots, X_n)\) such that

\[
M_n(\hat{\alpha}_n) \leq \min_{\alpha \in \mathcal{A}} M_n(\alpha) + c_n.
\]
∀ \( n \geq 1 \) and \( \alpha, \alpha' \in A \), there exist \( M'_n(\alpha) \), \( M''_n(\alpha) \) measurable functions depending on \( X_0, X_1, \ldots, X_n \) and on \( \alpha \) such that:

\[(A1) \forall \theta \in \Theta, \text{ there exists a unique } \alpha_0 = \alpha_0(\theta) \in A \text{ such that } M'_\theta(\alpha_0) = 0 \text{ where } M'_\theta(\alpha) = \lim_{n \to \infty} E_{\theta}[M'_n(\alpha)].

\[(A2) 0 < \inf_{\theta} m(\theta) \leq \sup_{\theta} m(\theta) < \infty \text{ where } m(\theta) = \lim_{n \to \infty} E_{\theta}[M''_n(\alpha_0)].

\[(A3) \forall n \geq 1, \exists r_n > 0 \text{ independent of } \theta \text{ such that } \sqrt{nr_n} \to 0 \text{ and } \sup_{\theta \in \Theta} P_{\theta}(|M'_n(\hat{\alpha}_n)| \geq r_n) = O(n^{-1/2}).

\[(A4) \text{ for } j = 1, 2, \exists \sigma_j(\cdot) \text{ such that } 0 < \inf_{\theta} \sigma_j(\theta) \leq \sup_{\theta} \sigma_j(\theta) < \infty, \text{ and some constant } B > 0 \text{ and integer } p_0 \text{ such that for all } n \geq p_0 \sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} |\sqrt{n} \sigma_j(\theta) M'_n(\alpha_0) \leq u | - \Gamma(u) | \leq B \sqrt{n} \sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} |\sqrt{n} \sigma_2(\theta)(M''_n(\alpha_0) - m(\theta)) \leq u | - \Gamma(u) | \leq B \sqrt{n} \).
Pfanzagl’s method revisited: Assumptions (1/3)

∀ \(n \geq 1\) and \(\alpha, \alpha' \in A\), there exist \(M'_n(\alpha)\), \(M''_n(\alpha)\) measurable functions depending on \(X_0, X_1, ..., X_n\) and on \(\alpha\) such that:

(A1) \(\forall \theta \in \Theta\), there exists a unique \(\alpha_0 = \alpha_0(\theta) \in A\) such that \(M'_\theta(\alpha_0) = 0\) where \(M'_\theta(\alpha) = \lim_{n \to \infty} E_\theta[M'_n(\alpha)]\).
Pfanzagl’s method revisited: Assumptions (1/3)

\(\forall \ n \geq 1 \text{ and } \alpha, \alpha' \in \mathcal{A}, \) there exist \(M_n'(\alpha), \ M_n''(\alpha)\) measurable functions depending on \(X_0, X_1, \ldots, X_n\) and on \(\alpha\) such that:

(A1) \(\forall \theta \in \Theta, \) there exists a unique \(\alpha_0 = \alpha_0(\theta) \in \mathcal{A}\) such that 
\(M_\theta'(\alpha_0) = 0\) where 
\(M_\theta'(\alpha) = \lim_{n \to \infty} \mathbb{E}_\theta[M_n'(\alpha)].\)

(A2) \(0 < \inf_\theta m(\theta) \leq \sup_\theta m(\theta) < \infty\) where 
\(m(\theta) = \lim_{n \to \infty} \mathbb{E}_\theta[M_n''(\alpha_0)].\)
Pfanzagl’s method revisited: Assumptions (1/3)

∀ \( n \geq 1 \) and \( \alpha, \alpha' \in A \), there exist \( M'_n(\alpha) \), \( M''_n(\alpha) \) measurable functions depending on \( X_0, X_1, \ldots, X_n \) and on \( \alpha \) such that:

(A1) \( \forall \theta \in \Theta \), there exists a unique \( \alpha_0 = \alpha_0(\theta) \in A \) such that \( M'_\theta(\alpha_0) = 0 \) where \( M'_\theta(\alpha) = \lim_{n \to \infty} \mathbb{E}_\theta[M'_n(\alpha)] \).

(A2) \( 0 < \inf_\theta m(\theta) \leq \sup_\theta m(\theta) < \infty \) where \( m(\theta) = \lim_{n \to \infty} \mathbb{E}_\theta[M''_n(\alpha_0)] \).

(A3) \( \forall n \geq 1, \exists r_n > 0 \) independent of \( \theta \) such that \( \sqrt{n}r_n \to 0 \) and

\[
\sup_{\theta \in \Theta} \mathbb{P}_\theta \left( |M'_n(\hat{\alpha}_n)| \geq r_n \right) = O(n^{-1/2}).
\]
∀ \( n \geq 1 \) and \( \alpha, \alpha' \in A \), there exist \( M'_n(\alpha) \), \( M''_n(\alpha) \) measurable functions depending on \( X_0, X_1, \ldots, X_n \) and on \( \alpha \) such that:

(A1) \( \forall \theta \in \Theta \), there exists a unique \( \alpha_0 = \alpha_0(\theta) \in A \) such that \( M'_\theta(\alpha_0) = 0 \) where \( M'_\theta(\alpha) = \lim_{n \to \infty} E_\theta[M'_n(\alpha)] \).

(A2) \( 0 < \inf_\theta m(\theta) \leq \sup_\theta m(\theta) < \infty \) where \( m(\theta) = \lim_{n \to \infty} E_\theta[M''_n(\alpha_0)] \).

(A3) \( \forall n \geq 1, \exists r_n > 0 \) independent of \( \theta \) such that \( \sqrt{n}r_n \to 0 \) and

\[
\sup_{\theta \in \Theta} \mathbb{P}_\theta \left( |M'_n(\hat{\alpha}_n)| \geq r_n \right) = O(n^{-1/2}).
\]

(A4) for \( j = 1, 2 \), \( \exists \sigma_j(\cdot) \) such that \( 0 < \inf_\theta \sigma_j(\theta) \leq \sup_\theta \sigma_j(\theta) < \infty \), and some constant \( B > 0 \) and integer \( P_0 \) such that for all \( n \geq P_0 \)

\[
\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_\theta \left\{ \frac{\sqrt{n}}{\sigma_1(\theta)} M'_n(\alpha_0) \leq u \right\} - \Gamma(u) \right| \leq \frac{B}{\sqrt{n}}
\]

\[
\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_\theta \left\{ \frac{\sqrt{n}}{\sigma_2(\theta)} (M''_n(\alpha_0) - m(\theta)) \leq u \right\} - \Gamma(u) \right| \leq \frac{B}{\sqrt{n}}.
\]
(A4') for \( n \geq P_0 \), \(|u| \leq 2\sqrt{\ln n} \), and \( \theta \in \Theta \), \( \exists \sigma_{n,u}(\theta) > 0 \) such that, with some constants \( A', B' > 0 \) which do not depend on such \( n, u, \theta \),

\[
|\sigma_{n,u}(\theta) - \sigma_1(\theta)| \leq A' |u| n^{-1/2},
\]

\[
\mathbb{P}_\theta \left\{ \frac{\sqrt{n}}{\sigma_{n,u}(\theta)} \left( M'_n(\alpha_0) + \frac{u \sigma_1(\theta)}{\sqrt{n} m(\theta)} (M''_n(\alpha_0) - m(\theta)) \right) \leq u \right\} - \Gamma(u) \leq \frac{B'}{\sqrt{n}}.
\]
(A4') for \( n \geq P_0, \) \( |u| \leq 2\sqrt{\ln n}, \) and \( \theta \in \Theta, \) \( \exists \sigma_{n,u}^2(\theta) > 0 \) such that, with some constants \( A', B' > 0 \) which do not depend on such \( n, u, \theta, \)

\[
|\sigma_{n,u}(\theta) - \sigma_1(\theta)| \leq A' |u| n^{-1/2},
\]

\[
\left| \mathbb{P}_\theta \left\{ \frac{\sqrt{n}}{\sigma_{n,u}(\theta)} \left( M'_n(\alpha_0) + \frac{u \sigma_1(\theta)}{\sqrt{n m(\theta)}} (M''_n(\alpha_0) - m(\theta)) \right) \right\} \leq u \right| - \Gamma(u) \right| \leq \frac{B'}{\sqrt{n}}.
\]

(A5) For any \( \alpha, \alpha' \in \mathcal{A}, \) let \( R_n(\alpha, \alpha') \) be defined by the equation

\[
M'_n(\alpha') = M'_n(\alpha) + [M''_n(\alpha) + R_n(\alpha, \alpha')] (\alpha' - \alpha).
\]

For each \( n \) there exist \( \omega_n \geq 0 \) independent of \( \theta \) and a real-valued measurable function \( W_n \) depending on \( X_0, X_1, \ldots, X_n \) but independent of \( \theta \) such that \( \omega_n \to 0 \) and \( \forall (\alpha, \alpha') \in \mathcal{A}^2, \)

\[
|R_n(\alpha, \alpha')| \leq \{|\alpha - \alpha'| + \omega_n\} W_n,
\]

and there exists a constant \( c_W > 0 \) such that

\[
\sup_{\theta \in \Theta} \mathbb{P}_\theta \{ c_W \leq W_n \} = O(n^{-1/2}).
\]
Assume that $\hat{\alpha}_n$ is uniformly consistent at some rate, that is there exists a sequence $\gamma_n \to 0$ such that

$$\sup_{\theta \in \Theta} \mathbb{P}_\theta \{ |\hat{\alpha}_n - \alpha_0| \geq d \} \leq \gamma_n,$$

where $d = \inf_{\theta \in \Theta} m(\theta)/8c_W$ with $c_W$ and $m(\theta)$ defined in (A5) and (A2) respectively.
Assume that $\hat{\alpha}_n$ is uniformly consistent at some rate, that is there exists a sequence $\gamma_n \to 0$ such that

$$\sup_{\theta \in \Theta} \mathbb{P}_\theta \{|\hat{\alpha}_n - \alpha_0| \geq d\} \leq \gamma_n,$$

where $d = \inf_{\theta \in \Theta} m(\theta)/8c_W$ with $c_W$ and $m(\theta)$ defined in (A5) and (A2) respectively.

**Comments:**
(A6) Assume that \( \hat{\alpha}_n \) is uniformly consistent at some rate, that is there exists a sequence \( \gamma_n \to 0 \) such that

\[
\sup_{\theta \in \Theta} \mathbb{P}_\theta \left\{ |\hat{\alpha}_n - \alpha_0| \geq d \right\} \leq \gamma_n,
\]

where \( d = \inf_{\theta \in \Theta} m(\theta)/8c_W \) with \( c_W \) and \( m(\theta) \) defined in (A5) and (A2) respectively.

**Comments:**
- The expectations \( \mathbb{E}_\theta [M_n'(\alpha)] \) and \( \mathbb{E}_\theta [M_n''(\alpha_0)] \) may depend on \( n \).
(A6) Assume that \( \hat{\alpha}_n \) is uniformly consistent at some rate, that is there exists a sequence \( \gamma_n \to 0 \) such that

\[
\sup_{\theta \in \Theta} \mathbb{P}_\theta \{ |\hat{\alpha}_n - \alpha_0| \geq d \} \leq \gamma_n,
\]

where \( d = \inf_{\theta \in \Theta} m(\theta)/8c_W \) with \( c_W \) and \( m(\theta) \) defined in (A5) and (A2) respectively.

**Comments:**
- The expectations \( \mathbb{E}_\theta [M'_n(\alpha)] \) and \( \mathbb{E}_\theta [M''_n(\alpha_0)] \) may depend on \( n \).
- Assumption (A3) ensures that the estimator (approximately) satisfies a kind of first order condition.
Pfanzagl’s method revisited: Assumptions (3/3)

(A6) Assume that \( \hat{\alpha}_n \) is uniformly consistent at some rate, that is there exists a sequence \( \gamma_n \to 0 \) such that

\[
\sup_{\theta \in \Theta} \mathbb{P}_\theta \{ |\hat{\alpha}_n - \alpha_0| \geq d \} \leq \gamma_n,
\]

where \( d = \inf_{\theta \in \Theta} m(\theta)/8c_W \) with \( c_W \) and \( m(\theta) \) defined in (A5) and (A2) respectively.

Comments:

- The expectations \( \mathbb{E}_\theta [M'_n(\alpha)] \) and \( \mathbb{E}_\theta [M''_n(\alpha_0)] \) may depend on \( n \).
- Assumption (A3) ensures that the estimator (approximately) satisfies a kind of first order condition.
- The Assumptions (A4) and (A4’) are the Berry-Esseen bounds for \( M'_n(\alpha_0), M''_n(\alpha_0) \), and for some of their linear combinations.
(A6) Assume that $\hat{\alpha}_n$ is uniformly consistent at some rate, that is there exists a sequence $\gamma_n \to 0$ such that

$$\sup_{\theta \in \Theta} \mathbb{P}_\theta \left\{ |\hat{\alpha}_n - \alpha_0| \geq d \right\} \leq \gamma_n,$$

where $d = \inf_{\theta \in \Theta} m(\theta)/8c_W$ with $c_W$ and $m(\theta)$ defined in (A5) and (A2) respectively.

**Comments:**

- The expectations $\mathbb{E}_\theta [M'_n(\alpha)]$ and $\mathbb{E}_\theta [M''_n(\alpha_0)]$ may depend on $n$.
- Assumption (A3) ensures that the estimator (approximately) satisfies a kind of first order condition.
- The Assumptions (A4) and (A4’) are the Berry-Esseen bounds for $M'_n(\alpha_0)$, $M''_n(\alpha_0)$, and for some of their linear combinations.
- The identity defining $R_n(\alpha, \alpha')$ in Assumption (A5) is guaranteed by Taylor expansion when the criterion $M_n(\alpha)$ is twice differentiable with respect to $\alpha$. 

**Theorem**

Under the conditions (A1-A6), \( \exists \ C > 0 \) such that \( \forall n \geq 1 \),

\[
\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta} \left\{ \frac{\sqrt{n}}{\tau(\theta)} (\hat{\alpha}_n - \alpha_0) \leq u \right\} - \Gamma(u) \right| \leq C \left( \frac{1}{\sqrt{n}} + \sqrt{nr_n} + \omega_n + \gamma_n \right)
\]

with \( \tau(\theta) := \sigma_1(\theta)/m(\theta) \).
Theorem

Under the conditions (A1-A6), \( \exists \ C > 0 \) such that \( \forall n \geq 1, \)

\[
\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta} \left\{ \frac{\sqrt{n}}{\tau(\theta)} (\hat{\alpha}_n - \alpha_0) \leq u \right\} - \Gamma(u) \right| \leq C \left( \frac{1}{\sqrt{n}} + \sqrt{n} r_n + \omega_n + \gamma_n \right)
\]

with \( \tau(\theta) := \sigma_1(\theta)/m(\theta). \)

Comments:

- To derive the Berry-Esseen bound of classical order \( O(n^{-1/2}) \), we need \( \gamma_n = O(n^{-1/2}), \ r_n = O(n^{-1}) \) and \( \omega_n = O(n^{-1/2}). \)

- Usually \( c_n \) in the definition of the \( M \)-estimator has to decrease at the rate \( n^{-3/2}. \)
1. Introduction

2. A general method for deriving Berry-Esseen bounds for $M$-estimators

3. A uniform Berry-Esseen statement with geometrically ergodic Markov chains

4. Application: Berry-Esseen bound for $M$-estimators with geometrically Markov chains

5. Example of application: AR(1) with ARCH(1) errors
Consider

- $E$ a measurable space with a countably generated $\sigma$-field $\mathcal{E}$
- $\Theta$ some general parameter space.
- $\{X_n\}_{n \geq 0}$ a Markov chain with state space $E$, transition kernels $\{Q_\theta(x, \cdot), \ x \in E\}, \ \theta \in \Theta,$ and initial distribution $\mu$.
- $\mu$ is either given (e.g. Dirac measure) or depends on $\theta$. Suppose $\mu(V) < \infty$ or $\sup_{\theta \in \Theta} \mu(V) < \infty$. 

Assumption (M).

$\forall \theta \in \Theta$, there exists a $Q_\theta$-invariant probability distribution $\pi_\theta$ and an unbounded $V : E \to [1, \infty)$ such that

(VG1) $\sup_{\theta \in \Theta} \pi_\theta(V) < \infty$

(VG2) For all $\gamma \in (0, 1]$, there exist real numbers $\kappa_\gamma < 1$ and $C_\gamma \geq 0$ such that we have, for each $\theta \in \Theta$ and for all $n \geq 1$ and $x \in E$

$$\sup_{f} |f| \leq V_\gamma |Q_n \theta f(x) - \pi_\theta(f)| \leq C_\gamma \kappa_\gamma V(x)_\gamma.$$ 

Remarks:

(VG1)-(VG2) $\forall \gamma \in (0, 1], \forall \theta \in \Theta$, $Q_\theta$ is $V_\gamma$-geometrically ergodic.

The constants $C_\gamma$ and $\kappa_\gamma$ do not depend on $\theta$. 

L. Hervé, J. Ledoux & V. Patilea ()

CFR 2010, Poitiers
Consider

- $E$ a measurable space with a countably generated $\sigma$-field $\mathcal{E}$
- $\Theta$ some general parameter space.
- $\{X_n\}_{n \geq 0}$ a Markov chain with state space $E$, transition kernels $\{Q_\theta(x, \cdot), \ x \in E\}$, $\theta \in \Theta$, and initial distribution $\mu$.
- $\mu$ is either given (e.g. Dirac measure) or depends on $\theta$. Suppose $\mu(V) < \infty$ or $\sup_{\theta \in \Theta} \mu(V) < \infty$.

Assumption ($\mathcal{M}$). $\forall \theta \in \Theta$, there exists a $Q_\theta$-invariant probability distribution $\pi_\theta$ and an unbounded $V : E \to [1, \infty)$ such that
Consider

- $E$ a measurable space with a countably generated $\sigma$-field $\mathcal{E}$
- $\Theta$ some general parameter space.
- $\{X_n\}_{n \geq 0}$ a Markov chain with state space $E$, transition kernels $\{Q_\theta(x, \cdot), \ x \in E\}$, $\theta \in \Theta$, and initial distribution $\mu$.
- $\mu$ is either given (e.g. Dirac measure) or depends on $\theta$. Suppose $\mu(V) < \infty$ or $\sup_{\theta \in \Theta} \mu(V) < \infty$.

Assumption (M). $\forall \theta \in \Theta$, there exists a $Q_\theta$-invariant probability distribution $\pi_\theta$ and an unbounded $V : E \to [1, \infty)$ such that

- (VG1) $\sup_{\theta \in \Theta} \pi_\theta(V) < \infty$

- (VG2) For all $\gamma \in (0, 1]$, there exist real numbers $\kappa_\gamma < 1$ and $C_\gamma \geq 0$ such that we have, for each $\theta \in \Theta$ and for all $n \geq 1$ and $x \in E$

$$\sup_{|f| \leq V^\gamma} |Q_\theta^n f(x) - \pi_\theta(f)| \leq C_\gamma \kappa_\gamma^n V(x)^\gamma.$$
Consider

- $E$ a measurable space with a countably generated $\sigma$-field $\mathcal{E}$
- $\Theta$ some general parameter space.
- $\{X_n\}_{n \geq 0}$ a Markov chain with state space $E$, transition kernels $\{Q_\theta(x, \cdot), \ x \in E\}$, $\theta \in \Theta$, and initial distribution $\mu$.
- $\mu$ is either given (e.g. Dirac measure) or depends on $\theta$. Suppose $\mu(V) < \infty$ or $\sup_{\theta \in \Theta} \mu(V) < \infty$.

Assumption (M). $\forall \theta \in \Theta$, there exists a $Q_\theta$-invariant probability distribution $\pi_\theta$ and an unbounded $V : E \to [1, \infty)$ such that

(VG1) $\sup_{\theta \in \Theta} \pi_\theta(V) < \infty$
(VG2) For all $\gamma \in (0, 1]$, there exist real numbers $\kappa_\gamma < 1$ and $C_\gamma \geq 0$ such that we have, for each $\theta \in \Theta$ and for all $n \geq 1$ and $x \in E$

$$\sup_{|f| \leq V^\gamma} |Q_\theta^n f(x) - \pi_\theta(f)| \leq C_\gamma \kappa_\gamma^n V(x)^\gamma.$$ 

Remarks:

- (VG1)-(VG2) $\Rightarrow \forall \gamma \in (0, 1]$, $\forall \theta \in \Theta$, $Q_\theta$ is $V^\gamma$-geometrically ergodic.
- The constants $C_\gamma$ and $\kappa_\gamma$ do not depend on $\theta$. 
Let $\alpha_0 = \alpha_0(\theta) \in \mathcal{A} \subset \mathbb{R}$ be the parameter of interest
Let $\alpha_0 = \alpha_0(\theta) \in A \subset \mathbb{R}$ be the parameter of interest

Let $\xi(\alpha, x, y)$ such that $\mathbb{E}_{\theta, \pi_{\theta}}[\xi(\alpha_0, X_0, X_1)] = 0$
Let $\alpha_0 = \alpha_0(\theta) \in A \subset \mathbb{R}$ be the parameter of interest

Let $\xi(\alpha, x, y)$ such that $\mathbb{E}_{\theta, \pi_{\theta}}[\xi(\alpha_0, X_0, X_1)] = 0$

Let $S_n(\alpha) = \sum_{k=1}^{n} \xi(\alpha, X_{k-1}, X_k)$
Let $\alpha_0 = \alpha_0(\theta) \in \mathcal{A} \subset \mathbb{R}$ be the parameter of interest.

Let $\xi(\alpha, x, y)$ such that $\mathbb{E}_{\theta, \pi_\theta} [\xi(\alpha_0, X_0, X_1)] = 0$.

Let $S_n(\alpha) = \sum_{k=1}^{n} \xi(\alpha, X_{k-1}, X_k)$.

We need to investigate the uniform Berry-Esseen property

$$
\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \mu} \left\{ \frac{S_n(\alpha_0)}{\sigma(\theta) \sqrt{n}} \leq u \right\} - \Gamma(u) \right| = O(n^{-1/2})
$$

where $\sigma^2(\theta)$ is the asymptotic variance.
Let $\alpha_0 = \alpha_0(\theta) \in A \subset \mathbb{R}$ be the parameter of interest.

Let $\xi(\alpha, x, y)$ such that $\mathbb{E}_{\theta, \pi_{\theta}} [\xi(\alpha_0, X_0, X_1)] = 0$.

Let $S_n(\alpha) = \sum_{k=1}^{n} \xi(\alpha, X_{k-1}, X_k)$.

We need to investigate the uniform Berry-Esseen property

$$
\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \mu} \left\{ \frac{S_n(\alpha_0)}{\sigma(\theta) \sqrt{n}} \leq u \right\} - \Gamma(u) \right| = O(n^{-1/2})
$$

where $\sigma^2(\theta)$ is the asymptotic variance.

The following condition will be required for $m_0 = 1, 2$ or $3$:

$\exists$ constants $m > m_0 \geq 1$ and $C_{\xi} > 0$ such that

$$
\forall \alpha \in A, \forall (x, y) \in E^2, \quad |\xi(\alpha, x, y)|^m \leq C_{\xi} (V(x) + V(y)). \quad (D_{m_0})
$$
Let $\alpha_0 = \alpha_0(\theta) \in A \subset \mathbb{R}$ be the parameter of interest

Let $\xi(\alpha, x, y)$ such that $\mathbb{E}_{\theta, \pi_{\theta}}[\xi(\alpha_0, X_0, X_1)] = 0$

Let $S_n(\alpha) = \sum_{k=1}^{n} \xi(\alpha, X_{k-1}, X_k)$

We need to investigate the uniform Berry-Esseen property

$$\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \mu} \left\{ \frac{S_n(\alpha_0)}{\sigma(\theta) \sqrt{n}} \leq u \right\} - \Gamma(u) \right| = O(n^{-1/2})$$

where $\sigma^2(\theta)$ is the asymptotic variance

The following condition will be required for $m_0 = 1, 2$ or 3:

$\exists$ constants $m > m_0 \geq 1$ and $C_\xi > 0$ such that

$$\forall \alpha \in A, \forall (x, y) \in E^2, \quad |\xi(\alpha, x, y)|^m \leq C_\xi (V(x) + V(y)). \quad (D_{m_0})$$

Condition $(D_{m_0})$ implies

$$\mathbb{E}_{\theta, \pi_{\theta}}[|\xi(\alpha, X_0, X_1)|^m] < \infty \quad (1)$$
Lemma

If $\xi$ is centered and satisfies condition $(D_{m_0})$ with $m_0 \in \mathbb{N}^*$, then there exists $\beta > 0$ such that $\forall \theta \in \Theta$, $\forall n \geq 1$, $\forall t \in [-\beta, \beta]$,

$$\mathbb{E}_{\theta,\mu}[e^{itS_n(\alpha_0)}] = \lambda_\theta(t)^n [1 + L_\theta(t)] + r_{\theta,n}(t),$$

- $\lambda_\theta(\cdot)$, $L_\theta(\cdot)$ and $r_{\theta,n}(\cdot)$ are $C^{m_0}([-\beta, \beta], \mathbb{C})$ functions
- $\lambda_\theta(0) = 1$, $\lambda'_\theta(0) = 0$, $L_\theta(0) = 0$ and $r_{\theta,n}(0) = 0$. 
Lemma

If $\xi$ is centered and satisfies condition $(D_{m_0})$ with $m_0 \in \mathbb{N}^*$, then there exists $\beta > 0$ such that $\forall \theta \in \Theta, \forall n \geq 1, \forall t \in [-\beta, \beta]$,

$$E_{\theta, \mu}[e^{itS_n(\alpha_0)}] = \lambda_\theta(t)^n [1 + L_\theta(t)] + r_{\theta, n}(t),$$

- $\lambda_\theta(\cdot), L_\theta(\cdot)$ and $r_{\theta, n}(\cdot)$ are $C^{m_0}([-\beta, \beta], \mathbb{C})$ functions
- $\lambda_\theta(0) = 1, \lambda'_\theta(0) = 0, L_\theta(0) = 0$ and $r_{\theta, n}(0) = 0$.

Furthermore, there exists $\rho \in (0, 1)$ such that for $\ell = 0, \ldots, m_0$:

$$G_\ell := \sup_{|t| \leq \beta, \theta \in \Theta, \ n \geq 1} \left\{ \rho^{-n} |r^{(\ell)}_{\theta, n}(t)| \right\} < \infty.$$ 

The constants $\beta, \rho, G_\ell$, and the following ones (for $\ell = 0, \ldots, m_0$)

$$E_\ell := \sup_{|t| \leq \beta, \theta \in \Theta} |\lambda^{(\ell)}_\theta(t)| < \infty, \quad F_\ell := \sup_{|t| \leq \beta, \theta \in \Theta} |L^{(\ell)}_\theta(t)| < \infty,$$

depend on $\xi$, but only through the constant $C_\xi$ of Condition $D_{m_0}$. 
**Proposition**

**Suppose**

- Assumption ($\mathcal{M}$) holds true
- the functional $\xi$ satisfies condition $(D_1)$ and is centered.

**Then,**

$$\sup_{\theta \in \Theta} \sup_{n \geq 1} \left| \mathbb{E}_{\theta, \mu} [S_n(\alpha_0)] \right| < \infty \quad \text{and} \quad \forall \theta \in \Theta, \lim_{n} \mathbb{E}_{\theta, \mu} [S_n(\alpha_0)/n] = 0.$$
Proposition

Suppose

- Assumption \((\mathcal{M})\) holds true
- the functional \(\xi\) satisfies condition \((D_1)\) and is centered.

Then,

\[
\sup_{\theta \in \Theta} \sup_{n \geq 1} \left| \mathbb{E}_{\theta, \mu} [S_n(\alpha_0)] \right| < \infty \quad \text{and} \quad \forall \theta \in \Theta, \lim_{n} \mathbb{E}_{\theta, \mu} [S_n(\alpha_0)/n] = 0.
\]

If in addition \(\xi\) satisfies condition \((D_2)\), then for each \(\forall \theta \in \Theta\),

\[
\sigma^2(\theta) := \lim_{n} \frac{\mathbb{E}_{\theta, \mu} [S_n(\alpha_0)^2]}{n} \in \mathbb{R} \quad \text{is well defined.}
\]
Proposition

Suppose

- Assumption \((M)\) holds true
- the functional \(\xi\) satisfies condition \((D_1)\) and is centered.

Then,

\[
\sup_{\theta \in \Theta} \sup_{n \geq 1} \left| \mathbb{E}_{\theta, \mu}[S_n(\alpha_0)] \right| < \infty \quad \text{and} \quad \forall \theta \in \Theta, \lim_n \mathbb{E}_{\theta, \mu}[S_n(\alpha_0)/n] = 0.
\]

If in addition \(\xi\) satisfies condition \((D_2)\), then for each \(\forall \theta \in \Theta,\)

\[
\sigma^2(\theta) := \lim_n \frac{\mathbb{E}_{\theta, \mu}[S_n(\alpha_0)^2]}{n} \in \mathbb{R} \quad \text{is well defined.}
\]

Furthermore, the function \(\sigma^2(\cdot)\) is bounded on \(\Theta\), and there exists a constant \(C > 0\) depending only on \(C_\xi\) and \(\mu(V)\) such that

\[
\forall \theta \in \Theta, \forall n \geq 1, \quad \left| n\sigma^2(\theta) - \mathbb{E}_{\theta, \mu}[S_n(\alpha_0)^2] \right| \leq C.
\]
Theorem

Assume that:

- condition \((M)\) holds true;
- the functional \(\xi\) is centered and satisfies condition \((D_3)\);
- \(\sigma_0^2 := \inf_{\theta \in \Theta} \sigma^2(\theta) > 0\).

Then there exists a constant \(B(\xi)\) such that

\[
\forall n \geq 1, \sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \mu}\left\{ \frac{S_n(\alpha_0)}{\sigma(\theta)\sqrt{n}} \leq u \right\} - \Gamma(u) \right| \leq \frac{B(\xi)}{\sqrt{n}}.
\]

Furthermore, the constant \(B(\xi)\) depends on the functional \(\xi\), but only through \(\sigma_0\) and the constant \(C_\xi\) of Condition \((D_3)\).
1. Introduction

2. A general method for deriving Berry-Esseen bounds for $M$–estimators

3. A uniform Berry-Esseen statement with geometrically ergodic Markov chains


5. Example of application: AR(1) with ARCH(1) errors
Consider a Markov chain satisfying the condition ($\mathcal{M}$).
Consider a Markov chain satisfying the condition \((M)\).

Let

\[
M_n(\alpha) = \frac{1}{n} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k)
\]

where \(\alpha \in \mathcal{A} \subset \mathbb{R}\) is the parameter of interest, \(F : \mathcal{A} \times E^2 \rightarrow \mathbb{R}\).
Consider a Markov chain satisfying the condition \((M)\).

Let

\[
M_n(\alpha) = \frac{1}{n} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k)
\]

where \(\alpha \in A \subset \mathbb{R}\) is the parameter of interest, \(F : A \times E^2 \rightarrow \mathbb{R}\).

Assume that \(F\) satisfies condition \((D_1)\) and let

\[
M_\theta(\alpha) = \lim_{n \to \infty} \mathbb{E}_{\theta,\mu}[M_n(\alpha)] = \mathbb{E}_{\theta,\pi_\theta}[F(\alpha, X_0, X_1)].
\]
Consider a Markov chain satisfying the condition (\(\mathcal{M}\)).

Let

\[
M_n(\alpha) = \frac{1}{n} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k)
\]

where \(\alpha \in \mathcal{A} \subset \mathbb{R}\) is the parameter of interest, \(F : \mathcal{A} \times E^2 \to \mathbb{R}\).

Assume that \(F\) satisfies condition \((D_1)\) and let

\[
M_\theta(\alpha) = \lim_{n \to \infty} \mathbb{E}_{\theta, \mu}[M_n(\alpha)] = \mathbb{E}_{\theta, \pi_\theta}[F(\alpha, X_0, X_1)].
\]

For each \(\theta \in \Theta\) there exists a unique \(\alpha_0 = \alpha_0(\theta) \in \mathcal{A}\), the “true” value of the parameter of interest, such that

\[
M_\theta(\alpha) > M_\theta(\alpha_0), \ \forall \alpha \neq \alpha_0.
\]
Consider a Markov chain satisfying the condition (\(\mathcal{M}\)).

Let

\[
M_n(\alpha) = \frac{1}{n} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k)
\]

where \(\alpha \in \mathcal{A} \subseteq \mathbb{R}\) is the parameter of interest, \(F : \mathcal{A} \times E^2 \rightarrow \mathbb{R}\).

Assume that \(F\) satisfies condition \((D_1)\) and let

\[
M_\theta(\alpha) = \lim_{n \to \infty} \mathbb{E}_{\theta,\mu}[M_n(\alpha)] = \mathbb{E}_{\theta,\pi}\theta[F(\alpha, X_0, X_1)].
\]

For each \(\theta \in \Theta\) there exists a unique \(\alpha_0 = \alpha_0(\theta) \in \mathcal{A}\), the “true” value of the parameter of interest, such that

\[
M_\theta(\alpha) > M_\theta(\alpha_0), \ \forall \alpha \neq \alpha_0.
\]

Consider a \(M\)- estimator \(\hat{\alpha}_n\), that is

\[
M_n(\hat{\alpha}_n) \leq \min_{\alpha \in \mathcal{A}} M_n(\alpha) + c_n,
\]

where \(c_n \downarrow 0\).
\[ \forall (x, y) \in E^2, \text{ the map } \alpha \mapsto F(\alpha, x, y) \text{ is twice continuously differentiable on } \mathcal{A}. \]
∀(x, y) ∈ E^2, the map α ↦ F(α, x, y) is twice continuously differentiable on A.

Let F′(·, ·, ·) and F″(·, ·, ·) be the first and second order partial derivatives w.r.t. α.
\( \forall (x, y) \in E^2 \), the map \( \alpha \mapsto F(\alpha, x, y) \) is twice continuously differentiable on \( \mathcal{A} \).

Let \( F'(\cdot, \cdot, \cdot) \) and \( F''(\cdot, \cdot, \cdot) \) be the first and second order partial derivatives w.r.t. \( \alpha \).

Let

\[
M'_n(\alpha) = \frac{1}{n} \sum_{k=1}^{n} F'(\alpha, X_{k-1}, X_k), \quad M''_n(\alpha) = \frac{1}{n} \sum_{k=1}^{n} F''(\alpha, X_{k-1}, X_k).
\]
Assumptions (1/2)

We impose sufficient conditions guaranteeing assumptions (A1)-(A6)

(V0) $F'$ and $F''$ satisfy Condition $(D_3)$;
We impose sufficient conditions guaranteeing assumptions (A1)-(A6)

(V0) $F'$ and $F''$ satisfy Condition $(D_3)$;

(V1) $\forall \theta \in \Theta, \quad E_{\theta, \pi_\theta}[F'(\alpha_0, X_0, X_1)] = 0$ and $\alpha_0 = \alpha_0(\theta)$ is unique with this property;
Assumptions (1/2)

We impose sufficient conditions guaranteeing assumptions (A1)-(A6)

(V0) $F'$ and $F''$ satisfy Condition $(D_3)$;
(V1) $\forall \theta \in \Theta$, $E_{\theta, \pi_\theta}[F'(\alpha_0, X_0, X_1)] = 0$ and $\alpha_0 = \alpha_0(\theta)$ is unique with this property;
(V2) $m(\theta) := E_{\theta, \pi_\theta}[F''(\alpha_0, X_0, X_1)]$ satisfies $\inf_{\theta \in \Theta} m(\theta) > 0;$
Assumptions (1/2)

We impose sufficient conditions guaranteeing assumptions (A1)-(A6)

(V0) \( F' \) and \( F'' \) satisfy Condition (D3);

(V1) \( \forall \theta \in \Theta, \ E_{\theta, \pi_\theta}[F'(\alpha_0, X_0, X_1)] = 0 \) and \( \alpha_0 = \alpha_0(\theta) \) is unique with this property;

(V2) \( m(\theta) := E_{\theta, \pi_\theta}[F''(\alpha_0, X_0, X_1)] \) satisfies \( \inf_{\theta \in \Theta} m(\theta) > 0 \);

(V3) \( M'_n(\hat{\alpha}_n) \) satisfies (A3), that is \( \forall n \geq 1 \ \exists r_n > 0 \) independent of \( \theta \) such that \( \sqrt{n}r_n \to 0 \) and \( \sup_{\theta \in \Theta} P_\theta (|M'_n(\hat{\alpha}_n)| \geq r_n) = O(n^{-1/2}) \).
Assumptions (1/2)

We impose sufficient conditions guaranteeing assumptions (A1)-(A6)

(V0) $F'$ and $F''$ satisfy Condition $(D_3)$;

(V1) $\forall \theta \in \Theta$, $\mathbb{E}_{\theta, \pi_\theta}[F'(\alpha_0, X_0, X_1)] = 0$ and $\alpha_0 = \alpha_0(\theta)$ is unique with this property;

(V2) $m(\theta) := \mathbb{E}_{\theta, \pi_\theta}[F''(\alpha_0, X_0, X_1)]$ satisfies $\inf_{\theta \in \Theta} m(\theta) > 0$;

(V3) $M'_n(\hat{\alpha}_n)$ satisfies (A3), that is $\forall n \geq 1 \exists r_n > 0$ independent of $\theta$ such that $\sqrt{n}r_n \to 0$ and $\sup_{\theta \in \Theta} \mathbb{P}_\theta(\|M'_n(\hat{\alpha}_n)\| \geq r_n) = O(n^{-1/2})$.

Define the asymptotic variances:

$$
\sigma_1^2(\theta) := \lim_n \frac{1}{n} \mathbb{E}_{\theta, \mu} \left[ \left( \sum_{k=1}^{n} F'(\alpha_0, X_{k-1}, X_k) \right)^2 \right],
$$

$$
\sigma_2^2(\theta) := \lim_n \frac{1}{n} \mathbb{E}_{\theta, \mu} \left[ \left( \sum_{k=1}^{n} F''(\alpha_0, X_{k-1}, X_k) - n m(\theta) \right)^2 \right].
$$
Assumptions (1/2)

We impose sufficient conditions guaranteeing assumptions (A1)-(A6)

(V0) \( F' \) and \( F'' \) satisfy Condition \((D_3)\);

(V1) \( \forall \theta \in \Theta, \ E_{\theta, \pi_\theta}[F'(\alpha_0, X_0, X_1)] = 0 \) and \( \alpha_0 = \alpha_0(\theta) \) is unique with this property;

(V2) \( m(\theta) := E_{\theta, \pi_\theta}[F''(\alpha_0, X_0, X_1)] \) satisfies \( \inf_{\theta \in \Theta} m(\theta) > 0 \);

(V3) \( M'_n(\hat{\alpha}_n) \) satisfies (A3), that is \( \forall n \geq 1 \ \exists r_n > 0 \) independent of \( \theta \) such that \( \sqrt{n} r_n \to 0 \) and \( \sup_{\theta \in \Theta} P_{\theta} (|M'_n(\hat{\alpha}_n)| \geq r_n) = O(n^{-1/2}) \).

Define the asymptotic variances:

\[
\sigma_1^2(\theta) := \lim_n \frac{1}{n} \ E_{\theta, \mu} \left[ \left( \sum_{k=1}^{n} F'(\alpha_0, X_{k-1}, X_k) \right)^2 \right]
\]

\[
\sigma_2^2(\theta) := \lim_n \frac{1}{n} \ E_{\theta, \mu} \left[ \left( \sum_{k=1}^{n} F''(\alpha_0, X_{k-1}, X_k) - n m(\theta) \right)^2 \right].
\]

Condition (V0) and Proposition 1 ensure \( \sup_{\theta \in \Theta} \sigma_j(\theta) < \infty \) for \( j = 1, 2 \).
Assumptions (2/2)

The following conditions are also assumed to hold

(V4) \( \inf_{\theta \in \Theta} \sigma_j(\theta) > 0 \) for \( j = 1, 2 \).
Assumptions (2/2)

The following conditions are also assumed to hold

(V4) \( \inf_{\theta \in \Theta} \sigma_j(\theta) > 0 \) for \( j = 1, 2 \).

(V5) There exist \( \eta \in (0, 1/2) \) and \( C > 0 \) such that
\[ \forall (\alpha, \alpha') \in \mathcal{A}^2, \forall (x, y) \in \mathcal{E}^2, \]
\[ |F''(\alpha, x, y) - F''(\alpha', x, y)| \leq C |\alpha - \alpha'| (V(x) + V(y))^{\eta}. \]
The following conditions are also assumed to hold

(V4) $\inf_{\theta \in \Theta} \sigma_j(\theta) > 0$ for $j = 1, 2$.

(V5) There exist $\eta \in (0, 1/2)$ and $C > 0$ such that
$$\forall (\alpha, \alpha') \in A^2, \forall (x, y) \in E^2,$$
$$\left| F''(\alpha, x, y) - F''(\alpha', x, y) \right| \leq C |\alpha - \alpha'| \left( V(x) + V(y) \right) \eta.$$

(V6) there exists a sequence $\gamma_n \to 0$ such that
$$\sup_{\theta \in \Theta} \mathbb{P}_\theta \left\{ |\hat{\alpha}_n - \alpha_0| \geq d \right\} \leq \gamma_n,$$

with $d = \inf_{\theta \in \Theta} m(\theta)/8\pi_\theta(V^\eta)$ and $\eta$ defined in (V5).
Theorem

Assume that the condition (M) holds true, that F satisfies condition (D₁), that conditions (V₀) to (V₆) are fulfilled. Let \( \tau(\theta) := \sigma₁(\theta)/m(\theta) \). Then there exists a positive constant C such that \( \forall n \geq 1 \),

\[
\sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \mu} \left\{ \frac{\sqrt{n}}{\tau(\theta)} (\hat{\alpha}_n - \alpha_0) \leq u \right\} - \Gamma(u) \right| \leq C \left( \frac{1}{\sqrt{n}} + \sqrt{n}r_n + \gamma_n \right).
\]
1. Introduction

2. A general method for deriving Berry-Esseen bounds for $M$–estimators

3. A uniform Berry-Esseen statement with geometrically ergodic Markov chains


5. Example of application: AR(1) with ARCH(1) errors
Condition (M) holds for $V(x) = (1 + |x|)^p$ with $p > 6$
Condition $\mathcal{M}$ holds for $V(x) = (1 + |x|)^p$ with $p > 6$

$F(\rho, X_{k-1}, X_k) = (X_k - \rho X_{k-1})^2$
Condition ($\mathcal{M}$) holds for $V(x) = (1 + |x|)^p$ with $p > 6$

$F(\rho, X_{k-1}, X_k) = (X_k - \rho X_{k-1})^2$ 

$F'(\rho, X_{k-1}, X_k) = -2X_{k-1}(X_k - \rho X_{k-1})$

$F''(\rho, X_{k-1}, X_k) = 2X_{k-1}^2$. 

By our results, there exists a positive constant $C$ such that 

$\forall n \geq 1, \sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \sqrt{n} \sigma_1(\theta) m(\theta) - 1 (\hat{\rho}_n - \rho_0) \leq u \right| - \Gamma(u) \leq C \sqrt{n}$. 

Estimation of the autoregressive coefficient $\rho_0$

- Condition ($\mathcal{M}$) holds for $V(x) = (1 + |x|)^p$ with $p > 6$

- $F(\rho, X_{k-1}, X_k) = (X_k - \rho X_{k-1})^2$

- $F'(\rho, X_{k-1}, X_k) = -2X_{k-1}(X_k - \rho X_{k-1})$
  $F''(\rho, X_{k-1}, X_k) = 2X_{k-1}^2$.

- By our results, there exists a positive constant $C$ such that

$$\forall n \geq 1, \sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \delta_0} \left\{ \frac{\sqrt{n}}{\sigma_1(\theta)m(\theta)^{-1}} (\hat{\rho}_n - \rho_0) \leq u \right\} - \Gamma(u) \right| \leq \frac{C}{\sqrt{n}}.$$
Condition $(\mathcal{M})$ holds for $V(x) = (1 + |x|)^p$ with $p > 12$
Condition (\(\mathcal{M}\)) holds for \(V(x) = (1 + |x|)^p\) with \(p > 12\)

Define

\[
Q_n(\gamma; r, v) = \frac{1}{n} \sum_{k=1}^{n} \eta_k^2(\gamma, r, v)
\]

where

\[
\eta_k(\gamma, r, v) = (X_k - rX_{k-1})^2 - v(1 - r^2 - \gamma) - \gamma X_{k-1}^2,
\]
Condition \((\mathcal{M})\) holds for \(V(x) = (1 + |x|)^p\) with \(p > 12\)

Define

\[
Q_n(\gamma; r, \nu) = \frac{1}{n} \sum_{k=1}^{n} \eta_k^2(\gamma, r, \nu)
\]

where

\[
\eta_k(\gamma, r, \nu) = (X_k - rX_{k-1})^2 - \nu(1 - r^2 - \gamma) - \gamma X_{k-1}^2,
\]

Let \(M_n(\gamma) = Q_n(\gamma; \hat{\rho}_n, \hat{\nu}_n^2)\), where \(\hat{\nu}_n^2 = n^{-1} \sum_{k=1}^{n} X_k^2\).
Condition $(\mathcal{M})$ holds for $V(x) = (1 + |x|^p)$ with $p > 12$

Define

$$Q_n(\gamma; r, v) = \frac{1}{n} \sum_{k=1}^{n} \eta_k^2(\gamma, r, v)$$

where

$$\eta_k(\gamma, r, v) = (X_k - rX_{k-1})^2 - v(1 - r^2 - \gamma) - \gamma X_{k-1}^2,$$

Let $M_n(\gamma) = Q_n(\gamma; \hat{\rho}_n, \hat{\tau}_n^2)$, where $\hat{\tau}_n^2 = n^{-1} \sum_{k=1}^{n} X_k^2$.

Define $\hat{\gamma}_n = \arg \min_{\gamma \in [m_\gamma, M_\gamma]} M_n(\gamma)$
Define

- $F'(\gamma, X_{k-1}, X_k) = 2(\tau_0^2 - X_{k-1}^2)\eta_k(\gamma, \rho_0, \tau_0^2)$
- $F''(\gamma, X_{k-1}, X_k) = 2(\tau_0^2 - X_{k-1}^2)^2$
- $M'_n(\gamma) = \frac{\partial Q_n}{\partial \gamma}(\gamma; \rho_0, \tau_0^2)$
- $M''_n(\gamma) = \frac{\partial^2 Q_n}{\partial \gamma^2}(\gamma; \rho_0, \tau_0^2)$. 
Define

- $F'(\gamma, X_{k-1}, X_k) = 2(\tau_0^2 - X_{k-1}^2)\eta_k(\gamma, \rho_0, \tau_0^2)$
- $F''(\gamma, X_{k-1}, X_k) = 2(\tau_0^2 - X_{k-1}^2)^2$
- $M'_n(\gamma) = \frac{\partial Q_n}{\partial \gamma}(\gamma; \rho_0, \tau_0^2)$
- $M''_n(\gamma) = \frac{\partial^2 Q_n}{\partial \gamma^2}(\gamma; \rho_0, \tau_0^2)$.

Check condition (V3) with $r_n = n^{-1}\log n$ and deduce that for some suitable $\tau(\theta)$

$$\forall n \geq 1, \quad \sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \delta_0} \left\{ \frac{\sqrt{n}}{\tau(\theta)} (\hat{\gamma}_n - \gamma_0) \leq u \right\} - \Gamma(u) \right| = O\left( \frac{\log n}{\sqrt{n}} \right).$$
Estimation of the coefficient $\gamma_0$ cont’d

- Define
  - $F'(\gamma, X_{k-1}, X_k) = 2(\tau_0^2 - X_{k-1}^2)\eta_k(\gamma, \rho_0, \tau_0^2)$
  - $F''(\gamma, X_{k-1}, X_k) = 2(\tau_0^2 - X_{k-1}^2)^2$
  - $M'_n(\gamma) = \partial Q_n/\partial \gamma(\gamma; \rho_0, \tau_0^2)$
  - $M''_n(\gamma) = \partial^2 Q_n/\partial \gamma^2(\gamma; \rho_0, \tau_0^2)$.

- Check condition (V3) with $r_n = n^{-1} \log n$ and deduce that for some suitable $\tau(\theta)$

$$\forall n \geq 1, \sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \delta_0} \left\{ \frac{\sqrt{n}}{\tau(\theta)} (\hat{\gamma}_n - \gamma_0) \leq u \right\} - \Gamma(u) \right| = O \left( \frac{\log n}{\sqrt{n}} \right).$$

- The log $n$ factor could be probably improved – open question...