A LLN for Weakly Stationary and Regular Processes

We have already established strong consistency and asymptotic normallity of the OLSE in the context of the AR(1) process wrt a strictly stationary and ergodic white noise process. But what would be the asymptotic behavior of the OLSE in the AR(1) if ergodicity is dropped? The following LLN provides a partial answer.

<u>Lemma</u> Suppose that $(x_t)_{t\in\mathbb{Z}}$ is weakly stationary and regular. Then $\frac{1}{T}\sum_{t=1}^{T} x_t \xrightarrow{P} E(x_0)$.

<u>Proof:</u> Let $\epsilon > 0$. Then due to Chebychev's inequality*

$$\mathbb{P}\left[\left|\frac{1}{T}\sum_{t=1}^{T} x_t - E(x_0)\right| > \epsilon\right] \leq \frac{E\left|\frac{1}{T}\sum_{t=1}^{T} x_t - E(x_0)\right|^2}{\epsilon^2}$$

$$= \frac{E\left[\frac{1}{T}\sum_{t=1}^{T} x_t - E(x_0)\right]^2}{\epsilon^2}$$

$$= \frac{E\left[\frac{1}{T}\sum_{t=1}^{T}(x_t - E(x_0))\right]^2}{\epsilon^2}$$

$$= \frac{1}{\epsilon^2 T^2}\left\{\sum_{t=1}^{T} E(x_t - E(x_0))^2 + \sum_{\substack{t,s=1\\ t\neq s}}^{T} E\left[(x_t - E(x_0))(x_s - E(x_0))\right]\right\}$$

$$\overset{\text{weak. stat.}}{=} \frac{1}{\epsilon^2 T^2}\left\{T\gamma_0 + 2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \ldots\right\}$$

$$= \frac{1}{\epsilon^2 T}\gamma_0 + \frac{2}{\epsilon^2 T^2}\sum_{t=1}^{T}(T-t)\gamma_t$$

The process is by assumption regular, hence $\gamma_T \to 0$ as $T\to\infty$. Thereby,

$$\frac{1}{T^2}\sum_{t=1}^{T-1}|T-t||\gamma_t| = \frac{1}{T^2}\sum_{t=1}^{T}|T-t||\gamma_t|$$

$$\leq \frac{1}{T^2}\sum_{t=1}^{T}|T-1||\gamma_t|$$

$$= \frac{T-1}{T^2}\sum_{t=1}^{T}|\gamma_t| \to 0 \quad \text{as} \quad T\to\infty$$

and thereby $\frac{1}{T^2}\sum_{t=1}^{T-1}(T-t)\gamma_t \to 0$ as $T \to +\infty$.

Also, it is obvious that $\frac{1}{T}\gamma_0 \to 0$ as $T \to \infty$.

So, since ϵ is chosen arbitrarily, $\lim_{T\to\infty} E\left[\frac{\left|\frac{1}{T}\sum_{t=1}^{T}x_t - E(x_0)\right|}{\epsilon}\right]^2 = 0$ for all $\epsilon > 0$

therefore, $\lim_{T\to\infty} \mathbb{P}\left[\left|\frac{1}{T}\sum_{t=1}^{T}x_t - E(x_0)\right| > \epsilon\right] = 0$ for all $\epsilon > 0$, which proves the lemma.

Remark. The previous proof implies that the convergence mode is essentially the stronger, compared to convergence in probability*, L^2 convergence (in mean square*)

*[Remember that, if we consider $\{X_n(\omega)\}_{n=1}^{\infty}$ and $X(\omega)$ random variables (or random vectors) defined on the probability space $(\Omega, f, \mathbb{P})$ then

- $X_n(\omega)$ converges in probability to $X(\omega)$ if and only if $\forall\, \epsilon > 0$
  $\lim_{n\to\infty} \mathbb{P}(\omega : \|X_n(\omega) - X(\omega)\| < \epsilon) = 1$
- Considering $E(\|X_n\|^r) < \infty$ for all n and $E(\|X\|^r) < \infty$,
  X_n converges in r'th mean (or in L^r) to X if and only if
  $\lim_{n\to\infty} E(\|X_n - X\|^r) = 0$
- Using the generalized Chebyshev's inequality we can show that convergence in r'th mean implies convergence in probability.
  Remember that for a random variable X the generalized Chebyshev's inequality says that for $r > 0$ if $E(|X|^r) < \infty$ then $\forall\, \epsilon > 0$
  $\mathbb{P}(|X| > \epsilon) \leq E(|X|^r)/\epsilon^r$]

Example. Consider the AR(1) process defined over a white noise process and $|b| < 1$ i.e. consider the recursion $y_t = f(y_{t-1}, \epsilon_t) = by_{t-1} + \epsilon_t$ $t \in \mathbb{Z}$ which we have already proved that it admits a unique solution $y = (y_t)_{t\in\mathbb{Z}}$ defined by $y_t = \sum_{j=0}^{\infty} b^j \epsilon_{t-j}$

Thereby, when $|b|<1$, y is a linear process for which absolute summability holds and as we have already seen, it is a weakly stationary and regular process.

According to the LLN for weakly stationary and regular processes stated in the aforementioned lemma:

$$\frac{1}{T}\sum_{t=1}^{T} y_t \xrightarrow{P} E(y_0)=0$$

Now, consider the asymptotic behavior of the OLSE for b in this context.

$$b_T = \frac{\sum_{t=1}^{T} y_t y_{t-1}}{\sum_{t=1}^{T} y_{t-1}^2} = \frac{\sum_{t=1}^{T}(b y_{t-1}+\epsilon_t)y_{t-1}}{\sum_{t=1}^{T} y_{t-1}^2} = \frac{b\sum_{t=1}^{T} y_{t-1}^2 + \sum_{t=1}^{T}\epsilon_t y_{t-1}}{\sum_{t=1}^{T} y_{t-1}^2} = b + \frac{\sum_{t=1}^{T}\epsilon_t y_{t-1}}{\sum_{t=1}^{T} y_{t-1}^2}$$

weak consistency of the OLSE could follow if $(\epsilon_t y_{t-1})_{t\in\mathbb{Z}}$ is weakly stationary and regular since then we could use the LLN presented, and $\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2$ appropriately converges.

Let's consider first the process $(\epsilon_t y_{t-1})_{t\in\mathbb{Z}}$.

1. $E(\epsilon_t y_{t-1}) = E\left(\epsilon_t \sum_{j=0}^{\infty} b^j \epsilon_{t-1-j}\right) = E\left(\sum_{j=0}^{\infty} b^j \epsilon_t \epsilon_{t-1-j}\right)$

$$= \sum_{j=0}^{\infty} b^j E(\epsilon_t \epsilon_{t-1-j}) \overset{WN}{=} 0$$

2. $E(\epsilon_t^2 y_{t-1}^2) = E\left[\epsilon_t^2 \left(\sum_{j=0}^{\infty} b^j \epsilon_{t-1-j}\right)^2\right] = E\left[\epsilon_t^2\left(\sum_{j=0}^{\infty} b^{2j}\epsilon_{t-1-j}^2 + \sum_{\substack{i,j=0\\ i\neq j}}^{\infty} b^i b^j \epsilon_{t-1-j}\epsilon_{t-1-i}\right)\right]$

$$= \sum_{j=0}^{\infty} b^{2j} E(\epsilon_t^2 \epsilon_{t-1-j}^2) + \sum_{\substack{i,j=0\\ i\neq j}}^{\infty} b^i b^j E(\epsilon_t^2 \epsilon_{t-1-j}\epsilon_{t-1-i}) \qquad (*)$$

but, no already stated assumptions, specify those two series.

3. $\gamma_k = E(\epsilon_t y_{t-1}\epsilon_{t-k} y_{t-k-1}) = E\left(\epsilon_t \sum_{j=0}^{\infty} b^j \epsilon_{t-1-j}\ \epsilon_{t-k}\sum_{j=0}^{\infty} b^j \epsilon_{t-1-k-j}\right)$

<u>Assumption:</u> $E(\epsilon_t^4) < +\infty \quad \forall\, t\in\mathbb{Z}$

Given that $E(\epsilon_t^4) < +\infty \ \forall t \in \mathbb{Z}$, due to the Cauchy-Schwarz inequality $E(\epsilon_t y_{t-1} \epsilon_{t-k} y_{t-k-1})$ exists.

Also,

$$E(\epsilon_t y_{t-1} \epsilon_{t-k} y_{t-k-1}) \overset{LIE}{=} E\left(E(\epsilon_t y_{t-1} \epsilon_{t-k} y_{t-k-1} / f_{t-1})\right) =$$

$$E(y_{t-1} \epsilon_{t-k} y_{t-k-1} \underbrace{E(\epsilon_t / f_{t-1})}) \quad \forall t \in \mathbb{Z},\ k>0, \text{ where } f_t = \sigma(\epsilon_{t-i},\ i \geq 0)$$

we know nothing about this conditional expectation

Assumption: $(\epsilon_t)_{t\in\mathbb{Z}}$ is m.d. process wrt $(f_t)_{t\in\mathbb{Z}}$

which implies that $\cdot$ (ϵ_t) is adopted to the filtration $\cdot$ $E|\epsilon_t| < +\infty \ \forall t \in \mathbb{Z}$

$\cdot$ $E(\epsilon_t / f_{t-1}) = 0 \ \forall t \in \mathbb{Z}$

(And given that $(\epsilon_t)_{t\in\mathbb{Z}}$ is a WN, $E|\epsilon_t^2| > 0 \ \forall t \in \mathbb{Z}$, (ϵ_t) is also a s.M.d process)

Therefore, $\gamma_k = E(\epsilon_t y_{t-1} \epsilon_{t-k} y_{t-k-1}) = 0 \quad \forall t \in \mathbb{Z},\ k>0$

which also implies that $(\epsilon_t y_{t-1})_{t\in\mathbb{Z}}$ is a regular process.

Now, we need to determine the two infinite series appearing in (*).

▶ If $E(\epsilon_t^4) < +\infty$, due to the Cauchy-Schwarz inequality $E(\epsilon_t^2 \epsilon_{t-1-j}^2) < +\infty \ \forall t \in \mathbb{Z},\ k \geq 0$

and $\sum_{j=0}^{\infty} \beta^{2j} E(\epsilon_t^2 \epsilon_{t-1-j}^2) \leq \sum_{j=0}^{\infty} \beta^{2j} \sup_{j\geq 0} E(\epsilon_t^2 \epsilon_{t-1-j}^2)$ given that $\sup_{j\geq 0} E(\epsilon_t^2 \epsilon_{t-1-j}^2) < +\infty$

$$= \sup_{j\geq 0} E(\epsilon_t^2 \epsilon_{t-1-j}^2) \sum_{j=0}^{\infty} \beta^{2j}$$

$$= r \frac{1}{1-\beta^2} < +\infty \quad \text{where } r = \sup_{j\geq 0} E(\epsilon_t^2 \epsilon_{t-1-j}^2)$$

hence due to monotonicity $\sum_{j=0}^{\infty} \beta^{2j} E(\epsilon_t^2 \epsilon_{t-1-j}^2)$ exists.

Assumption: $(\epsilon_t)_{t\in\mathbb{Z}}$ is strictly stationary

then $\sum_{j=0}^{\infty} \beta^{2j} E(\epsilon_t^2 \epsilon_{t-1-j}^2) < +\infty$ is independent of t

▸ If $E(\epsilon_t^4) < +\infty \ \forall t \in \mathbb{Z}$, due to the Cauchy-Schwarz inequality
$E(\epsilon_t^2 \epsilon_{t-1-j} \epsilon_{t-1-i}) < +\infty \ \forall t \in \mathbb{Z}, \forall i,j \geq 0, i \neq j$
Also,
$$E(\epsilon_t^2 \epsilon_{t-1-j} \epsilon_{t-1-i}) \overset{LIE}{=} E\left(E\left(\epsilon_t^2 \epsilon_{t-1-j} \epsilon_{t-1-i} / f_{t-\min(j,i)-2}\right)\right) =$$
$$E\left(\epsilon_{t-1-\max(j,i)} \underbrace{E\left(\epsilon_t^2 \epsilon_{t-1-\min(j,i)} / f_{t-\min(j,i)-2}\right)}_{\text{we have no information about this conditional expectation.}}\right)$$
Assumption: $E(\epsilon_t^2 \epsilon_{t-1-j} / f_{t-j-2}) = 0 \ \forall t \in \mathbb{Z}, \forall j \geq 0$

then $E(\epsilon_t^2 \epsilon_{t-1-j} \epsilon_{t-1-i}) = 0$.

Therefore, condition 2 is formed as:
$(*)$: $E(\epsilon_t^2 y_{t-1}^2) = \sum_{j=0}^{\infty} \phi^{2j} E(\epsilon_t^2 \epsilon_{t-1-j}^2) < +\infty$ independent of t.

Hence, we have proved that $(\epsilon_t y_{t-1})_{t \in \mathbb{Z}}$ is a weakly stationary and regular process.
Wrapping up all the previous assumptions, consider the following.

<u>Assumption [1].</u> $(\epsilon_t)_{t \in \mathbb{Z}}$ is strictly stationary m.d. process wrt $(f_t)_{t \in \mathbb{Z}}$ with $f_t = \sigma(\epsilon_{t-i}, i \geq 0)$, such that $E(\epsilon_0^4) < +\infty$ and $(v_k)_{k \in \mathbb{N}}$ defined by $v_k = E(\epsilon_0^2 \epsilon_{-1-k}^2)$ is bounded and $E(\epsilon_t^2 \epsilon_{t-1-j} / f_{t-j-2}) = 0 \ \forall t \in \mathbb{Z}, \forall j \geq 0$. Finally, $v_k > 0$ for at least one $k \geq 0$.

So, essentially we have proved the following lemma.
<u>Lemma:</u> Under Assumption [1] consider the AR(1) process defined on $(\epsilon_t)_{t \in \mathbb{Z}}$ for $|\phi| < 1$. Then the process $(\epsilon_t y_{t-1})_{t \in \mathbb{Z}}$ is a white noise process with variance equal to $\sum_{j=0}^{\infty} \phi^{2j} E(\epsilon_0^2 \epsilon_{-1-j}^2)$.

Hence, by the LLN for weakly stationary and regular processes

$$\frac{1}{T}\sum_{t=1}^{T} \epsilon_t y_{t-1} \xrightarrow{P} E(\epsilon_0 y_{-1}) = 0$$

Consider now the asymptotic behavior of (y_{t-1}^2).
Remember that $y_{t-1} = \sum_{j=0}^{\infty} b^j \epsilon_{t-1-j}$, thus since according to Assumption [1] $(\epsilon_t)_{t\in\mathbb{Z}}$ is strictly stationary, $(y_{t-1})_{t\in\mathbb{Z}}$ is also strictly stationary.
Thereby, $(y_{t-1}^2)_{t\in\mathbb{Z}}$ is also a strictly stationary process.
By Doob's LLN, $\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 \longrightarrow E(y_0^2/\mathcal{I})$ $\mathbb{P}$ a.s.

therefore, it converges also in probability to the same limit. (Remember that almost sure convergence implies convergence in probability).
Suppose that $E(y_0^2/\mathcal{I}) > 0$ $\mathbb{P}$ a.s.
Then, due to the previous derivations and lemmata

$$\begin{pmatrix} 1/T \sum_{t=1}^{T} \epsilon_t y_{t-1} \\ 1/T \sum_{t=1}^{T} y_{t-1}^2 \end{pmatrix} \xrightarrow{P} \begin{pmatrix} 0 \\ E(y_0^2/\mathcal{I}) \end{pmatrix}$$

and since $E(y_0^2/\mathcal{I}) > 0$ $\mathbb{P}$ a.s., due to the CMT:

$$b_T = b + \frac{\frac{1}{T}\sum_{t=1}^{T} \epsilon_t y_{t-1}}{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2} \xrightarrow{P} b + \frac{0}{E(y_0^2/\mathcal{I})} = b$$

Thereby, we have essentially proven that:
Theorem. Under Assumption [1], for the AR(1) process defined on $(\epsilon_t)_{t\in\mathbb{Z}}$, for $|b|<1$ and if $E(y_0/\mathcal{I}) > 0$ $\mathbb{P}$ a.s. then the OLSE is weakly consistent.

Hence, by "sacrificing" ergodicity we obtain a weaker form of consistency via the use of assumptions that somehow "control" the dependence of the structuring white noise process, in the framework constructed above, and stricter moment existence conditions.