Chapter 7

Mixingales

In this section we prove some of the results stated in the previous sections using mixingales.

We first define a mixingale, noting that the definition we give is not the most general definition.

Definition 7.0.1 (Mixingale) *Let* $\mathcal{F}_t = \sigma(X_t, X_{t-1}, \ldots)$, $\{X_t\}$ *is called a mixingale if it satisfies*

$$\rho_{t,k} = \left\{ \mathbb{E}\left(\mathbb{E}(X_t|\mathcal{F}_{t-k}) - \mathbb{E}(X_t) \right)^2 \right\}^{1/2},$$

where $\rho_{t,k} \to 0$ *as* $k \to \infty$. *We note if* $\{X_t\}$ *is a stationary process then* $\rho_{t,k} = \rho_k$.

Lemma 7.0.3 *Suppose* $\{X_t\}$ *is a mixingale. Then* $\{X_t\}$ *almost surely satisfies the decomposition*

$$X_t = \sum_{j=0}^{\infty} \left\{ \mathbb{E}(X_t|\mathcal{F}_{t-j-1}) - \mathbb{E}(X_t|\mathcal{F}_{t-j-1}) \right\}. \tag{7.1}$$

PROOF. We first note that by using a telescoping argument that

$$X_t - \mathbb{E}(X_t) \;\; = \;\; \sum_{k=0}^{m} \left\{ \mathbb{E}(X_t|\mathcal{F}_{t-k}) - \mathbb{E}(X_t|\mathcal{F}_{t-k-1}) \right\} + \left\{ \mathbb{E}(X_t|\mathcal{F}_{t-m-1}) - \mathbb{E}(X_t) \right\}.$$

By definition of a martingale $\mathbb{E}\left(\mathbb{E}(X_t|\mathcal{F}_{t-m-1}) - \mathbb{E}(X_t) \right)^2 \to 0$ as $k \to \infty$, hence the remainder term in the above expansion becomes negligable as $m \to \infty$ and we have almost surely

$$X_t - \mathbb{E}(X_t)$$
$$= \;\; \sum_{k=0}^{\infty} \left\{ \mathbb{E}(X_t|\mathcal{F}_{t-k}) - \mathbb{E}(X_t|\mathcal{F}_{t-k-1}) \right\}.$$

Thus giving the required result. □

We observe that (7.1) resembles the Wold decomposition. The difference is that the Wolds decomposition decomposes a stationary process into elements which are the errors in the best

linear predictors. Whereas the result above decomposes a process into sums of martingale differences.

It can be shown that functions of several ARCH-type processes are mixingales (where $\rho_{t,k} \leq K\rho^k$ ($rho < 1$)), and Subba Rao (2006) and Dahlhaus and Subba Rao (2007) used these properties to obtain the rate of convergence for various types of ARCH parameter estimators. In a series of papers, Wei Biao Wu considered properties of a general class of stationary processes which satisfied Definition 7.0.1, where $\sum_{k=1}^{\infty} \rho_k < \infty$.

In Section 7.2 we use the mixingale property to prove Theorem 6.1.3. This is a simple illustration of how useful mixingales can be. In the following section we give a result on the rate of convergence of some random variables.

7.1 Obtaining almost sure rates of convergence for some sums

The following lemma is a simple variant on a result proved in Móricz (1976), Theorem 6.

Lemma 7.1.1 *Let $\{\mathcal{S}_T\}$ be a random sequence where $\mathbb{E}(\sup_{1 \leq t \leq T} |\mathcal{S}_t|^2) \leq \phi(T)$ and $\{phi(t)\}$ is a monotonically increasing sequence where $\phi(2^j)/\phi(2^{j-1}) \leq K < \infty$ for all j. Then we have almost surely*

$$\frac{1}{T}\mathcal{S}_T = O\big(\frac{\sqrt{\phi(T)(\log T)(\log\log T)^{1+\delta}}}{T}\big).$$

PROOF. The idea behind the proof is to that we find a subsequence of the natural numbers and define a random variables on this subsequence. This random variable, should 'dominate' (in some sense) S_T. We then obtain a rate of convergence for the subsequence (you will see that for the subsequence its quite easy by using the Borel-Cantelli lemma), which, due to the dominance, can be transfered over to S_T. We make this argument precise below.

Define the sequence $V_j = \sup_{t \leq 2^j} |S_t|$. Using Chebyshev's inequality we have

$$P(V_j > \varepsilon) \leq \frac{\phi(2^j)}{\varepsilon}.$$

Let $\varepsilon(t) = \sqrt{\phi(t)(\log\log t)^{1+\delta}\log t}$. It is clear that

$$\sum_{j=1}^{\infty} P(V_j > \varepsilon(2^j)) \leq \sum_{j=1}^{\infty} \frac{C\phi(2^j)}{\phi(2^j)(\log j)^{1+\delta}j} < \infty,$$

where C is a finite constant. Now by Borel Cantelli, this means that almost surely $V_j \leq \varepsilon(2^j)$. Let us now return to the orginal sequence S_T. Suppose $2^{j-1} \leq T \leq 2^j$, then by definition of V_j we have

$$\frac{S_T}{\varepsilon(T)} \leq \frac{V_j}{\varepsilon(2^{j-1})} \overset{a.s}{\leq} \frac{\varepsilon(2^j)}{\varepsilon(2^{j-1})} < \infty$$

under the stated assumptions. Therefore almost surely we have $S_T = O(\varepsilon(T))$, which gives us the required result. $\qquad\square$

We observe that the above result resembles the law of iterated logarithms. The above result is very simple and nice way of obtaining an almost sure rate of convergence. The main problem is obtaining bounds for $\mathbb{E}(\sup_{1 \leq t \leq T} |\mathcal{S}_t|^2)$. There is on exception to this, when $\mathcal{S}_t$ is the sum of martingale differences then one can simply apply Doob's inequality, where $\mathbb{E}(\sup_{1 \leq t \leq T} |\mathcal{S}_t|^2) \leq \mathbb{E}(|\mathcal{S}_T|^2)$. In the case that S_T is not the sum of martingale differences then its not so straightforward. However if we can show that S_T is the sum of mixingales then with some modifications a bound for $\mathbb{E}(\sup_{1 \leq t \leq T} |\mathcal{S}_t|^2)$ can be obtained. We will use this result in the section below.

7.2 Proof of Theorem 6.1.3

We summarise Theorem 6.1.3 below.

Theorem 1 *Let us suppose that $\{X_t\}$ has an ARMA representation where the roots of the characteristic polynomials $\phi(z)$ and $\theta(z)$ lie are greater than $1 + \delta$. Then*

(i)

$$\frac{1}{n} \sum_{t=r+1}^{n} \varepsilon_t X_{t-r} = O(\sqrt{\frac{(\log \log n)^{1+\gamma} \log n}{n}}) \tag{7.2}$$

(ii)

$$\frac{1}{n} \sum_{t=\max(i,j)}^{n} X_{t-i} X_{t-j} = O(\sqrt{\frac{(\log \log n)^{1+\gamma} \log n}{n}}). \tag{7.3}$$

for any $\gamma > 0$.

By using Lemma **??**, and that $\sum_{t=r+1}^{n} \varepsilon_t X_{t-r}$ is the sum of martingale differences, we prove Theorem 6.1.3(i) below.

PROOF of Theorem 6.1.3. We first observe that $\{\varepsilon_t X_{t-r}\}$ are martingale differences, hence we can use Doob's inequality to give $\mathbb{E}(\sup_{r+1 \leq s \leq T} (\sum_{t=r+1}^{s} \varepsilon_t X_{t-r})^2) \leq (T-r)\mathbb{E}(\varepsilon_t^2)\mathbb{E}(X_t^2)$. Now we can apply Lemma **??** to obtain the result. □

We now show that

$$\frac{1}{T} \sum_{t=\max(i,j)}^{T} X_{t-i} X_{t-j} = O(\sqrt{\frac{(\log \log T)^{1+\delta} \log T}{T}}).$$

However the proof is more complex, since $\{X_{t-i} X_{t-j}\}$ are not martingale differences and we cannot directly use Doob's inequality. However by showing that $\{X_{t-i} X_{t-j}\}$ is a mixingale we can still show the result.

To prove the result let $\mathcal{F}_t = \sigma(X_t, X_{t-1}, \ldots)$ and $\mathcal{G}_t = \sigma(X_{t-i} X_{t-j}, X_{t-1-i} X_{t-j-i}, \ldots)$. We observe that if $i > j$, then $\mathcal{G}_t \subset \mathcal{F}_{t-i}$.

Lemma 7.2.1 *Let $\mathcal{F}_t = \sigma(X_t, X_{t-1}, \ldots)$ and suppose X_t comes from an ARMA process, where the roots are greater than $1 + \delta$. Then if $\mathbb{E}(\varepsilon_t^4) < \infty$ we have*

$$\mathbb{E}\big(\mathbb{E}(X_{t-i}X_{t-j}|\mathcal{F}_{t-\min(i,j)-k}) - \mathbb{E}(X_{t-i}X_{t-j})\big)^2 \le C\rho^k.$$

PROOF. By expanding X_t as an MA(∞) process we have

$$\mathbb{E}(X_{t-i}X_{t-j}|\mathcal{F}_{t-\min(i,j)-k}) - \mathbb{E}(X_{t-i}X_{t-j})$$

$$= \sum_{j_1,j_2=0}^{\infty} a_{j_1} a_{j_2} \big\{\mathbb{E}(\varepsilon_{t-i-j_1}\varepsilon_{t-j-j_2}|\mathcal{F}_{t-k-\min(i,j)}) - \mathbb{E}(\varepsilon_{t-i-j_1}\varepsilon_{t-j-j_2})\big\}.$$

Now in the case that $t - i - j_1 > t - k - \min(i,j)$ and $t - j - j_2 > t - k - \min(i,j)$, $\mathbb{E}(\varepsilon_{t-i-j_1}\varepsilon_{t-j-j_2}|\mathcal{F}_{t-k-\min(i,j)}) = \mathbb{E}(\varepsilon_{t-i-j_1}\varepsilon_{t-j-j_2})$. Now by considering when $t - i - j_1 \le t - k - \min(i,j)$ or $t - j - j_2 \le t - k - \min(i,j)$ we have have the result. $\square$

Lemma 7.2.2 *Suppose $\{X_t\}$ comes from an ARMA process. Then*

(i) *The sequence $\{X_{t-i}X_{t-j}\}_t$ satisfies the mixingale property*

$$\mathbb{E}\big(\mathbb{E}(X_{t-i}X_{t-j}|\mathcal{F}_{t-\min(i,j)-k}) - \mathbb{E}(X_{t-i}X_{t-j}|\mathcal{F}_{t-k-1})\big)^2 \le K\rho^k, \tag{7.4}$$

and almost surely we can write $X_{t-i}X_{t-j}$ as

$$X_{t-i}X_{t-j} - \mathbb{E}(X_{t-i}X_{t-j}) = \sum_{k=0}^{\infty}\sum_{t=\min(i,j)}^{n} V_{t,k} \tag{7.5}$$

where $V_{t,k} = \mathbb{E}(X_{t-i}X_{t-j}|\mathcal{F}_{t-k-\min(i,j)}) - \mathbb{E}(X_{t-i}X_{t-j}|\mathcal{F}_{t-k-\min(i,j)-1})$, are martingale differences.

(ii) *Furthermore $\mathbb{E}(V_{t,k}^2) \le K\rho^k$ and*

$$\mathbb{E}\big\{\sup_{\min(i,j)\le s\le n}\big(\sum_{t=\min(i,j)}^{s}\{X_{t-i}X_{t-j} - \mathbb{E}(X_{t-i}X_{t-j})\}\big)^2\big\} \le Kn, \tag{7.6}$$

where K is some finite constant.

PROOF. To prove (i) we note that by using Lemma 7.2.1 we have (7.4). To prove (7.5) we use the same telescoping argument used to prove Lemma 7.0.3.

To prove (ii) we use the above expansion to give

$$\mathbb{E}\big\{\sup_{\min(i,j)\le s\le n}\big(\sum_{t=\min(i,j)}^{s}\{X_{t-i}X_{t-j} - \mathbb{E}(X_{t-i}X_{t-j})\}\big)^2\big\} \tag{7.7}$$

$$= \mathbb{E}\big\{\sup_{\min(i,j)\le s\le n}\big(\sum_{k=0}^{\infty}\sum_{t=\min(i,j)}^{s} V_{t,k}\big)^2\big\}$$

$$= \mathbb{E}\big\{\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sup_{\min(i,j)\le s\le n}\big|\sum_{t=\min(i,j)}^{s} V_{t,k_1}\big| \times \big|\sum_{t=\min(i,j)}^{s} V_{t,k_2}\big|\big\}$$

$$= \Big(\sum_{k=0}^{\infty}\Big\{\mathbb{E}\Big(\sup_{\min(i,j)\le s\le n}\big|\sum_{t=\min(i,j)}^{s} V_{t,k_1}\big|^2\Big)\Big\}^{1/2}\Big)^2$$

77

Now we see that $\{V_{t,k}\}_t = \{\mathbb{E}(X_{t-i}X_{t-j}|\mathcal{F}_{t-k-\min(i,j)}) - \mathbb{E}(X_{t-i}X_{t-j}|\mathcal{F}_{t-k-\min(i,j)-1})\}_t$, therefore $\{V_{t,k}\}_t$ are also martingale differences. Hence we can apply Doob's inequality to $\mathbb{E}\{\sup_{\min(i,j)\leq s\leq n}\left(\sum_{t=\min(i,j)}^{s}V_{t,k}\right)$ and by using (7.4) we have

$$\mathbb{E}\{\sup_{\min(i,j)\leq s\leq n}\left(\sum_{t=\min(i,j)}^{s}V_{t,k}\right)^2\} \leq \mathbb{E}\left(\sum_{t=\min(i,j)}^{n}V_{t,k}\right)^2 = \sum_{t=\min(i,j)}^{n}\mathbb{E}(V_{t,k}^2) \leq K\cdot n\rho^k.$$

Therefore now by using (7.7) we have

$$\mathbb{E}\{\sup_{\min(i,j)\leq s\leq n}\left(\sum_{t=\min(i,j)}^{s}\{X_{t-i}X_{t-j} - \mathbb{E}(X_{t-i}X_{t-j})\}\right)^2\} \leq Kn.$$

Thus giving (7.6). $\qquad\square$

We now use the above to prove Theorem 6.1.3(ii).

PROOF of Theorem 6.1.3(ii). To prove the result we use (7.6) and Lemma 7.1.1. $\quad\square$