Martingale approximations for continuous-time and discrete-time stationary Markov processes

Hajo Holzmann* 1

Institut für Mathematische Stochastik, Georg-August-Universität Göttingen, Maschmühlenweg 8-10, D-37073 Göttingen, Germany

Received 1 July 2004; received in revised form 14 March 2005; accepted 1 April 2005
Available online 27 April 2005

Abstract

We show that the method of Kipnis and Varadhan [Comm. Math. Phys. 104 (1986) 1–19] to construct a Martingale approximation to an additive functional of a stationary ergodic Markov process via the resolvent is universal in the sense that a martingale approximation exists if and only if the resolvent representation converges. A sufficient condition for the existence of a martingale approximation is also given. As examples we discuss moving average processes and processes with normal generator.

MSC: primary 60F05; 60G44; secondary 60J25; 60J35

Keywords: Markov process; Central limit theorem; Martingale; Moving average process; Normal operator

1. Introduction

The central limit theorem (CLT) for additive functionals of stationary, ergodic Markov chains has been studied intensively during the last decades. A basic approach for proving the CLT, initiated by Gordin and Lifšic [13] and afterwards...
pursued by several authors, is to construct a martingale approximation to the partial sums. These are decomposed into a sum of a martingale with stationary increments and a remainder term. After showing that the remainder term is negligible in some suitable sense, asymptotic normality follows from a martingale CLT. In this note we will focus on the case where the remainder is negligible in mean-square. Let \((X_n)_{n \geq 0}\) be a stationary ergodic (discrete-time) Markov chain with state space \((X, \mathcal{B})\), transition operator \(Q\) and stationary initial distribution \(\mu\). We denote by \(\| \cdot \|\) and \(\langle \cdot, \cdot \rangle\) the norm and the inner product of \(L_2(\mu)\), respectively, and by \(L_2^0(\mu)\) the subspace with \(\int f \, d\mu = 0\). For a fixed function \(f \in L_2^0(\mu)\) let \(S_0 = 0\) and

\[
S_n(f) = f(X_1) + \cdots + f(X_n), \quad n \geq 1.
\]

**Definition 1.1.** We say that there is a martingale approximation to \(S_n(f)\) if there exist two sequences of random variables \((M_n)_{n \geq 1}\) and \((A_n)_{n \geq 1}\) such that

1. \(S_n(f) = M_n + A_n, \quad n \geq 1\),
2. \((M_n)_{n \geq 1}\) is a square-integrable martingale with stationary increments with respect to \(\mathcal{F}_n = \sigma(X_0, \ldots, X_n)\),
3. \(E(A_n)^2/n \to 0, \quad n \to \infty\).

Notice that if there exists a martingale approximation to \(S_n(f)\), the processes \((M_n)_{n \geq 1}\) and \((A_n)_{n \geq 1}\) are uniquely determined a.s.. Given a martingale approximation to \(S_n(f)\), from the CLT for martingales with stationary, ergodic increments due to Billingsley [2] and Ibragimov [16] it follows that:

\[
\frac{S_n(f)}{\sqrt{n}} \overset{n \to \infty}{\Rightarrow} \mathcal{N}(0, \sigma^2(f)),
\]

where the asymptotic variance satisfies

\[
\sigma^2(f) = EM_1^2 = \lim_{n \to \infty} ES_n(f)^2 / n.
\]

A martingale approximation in the above sense (with additional properties in several cases) was constructed by Derriennic and Lin [9–11], Gordin and Holzmann [12], Gordin and Lifšic [13,14], Kipnis and Varadhan [17], Maxwell and Woodroofe [18] and Woodroofe [20] under suitable conditions on the function \(f\) and in some cases on the Markov operator \(Q\). Wu and Woodroofe [21] also investigated necessary and sufficient conditions for existence of (a different notion of) martingale approximations. For a comparison with their results see Remark 2.4.

In this note we will mainly consider continuous-time Markov processes. Let \((X_t)_{t \geq 0}\) be a stationary ergodic Markov process, defined on a probability space \((\Omega, \mathcal{F}, P)\), with state space \((X, \mathcal{B})\), transition probability function \(p(t, x, dy)\) and stationary initial distribution \(\mu\). We assume that the contraction semigroup

\[
T_tf(x) = \int_X f(y) p(t, x, dy), \quad f \in L_2(\mu)
\]

is strongly continuous (on \(L_2(\mu)\)). Let \((\mathcal{F}_t)_{t \geq 0}\) be a filtration in \((\Omega, \mathcal{F}, P)\) such that \((X_t)_{t \geq 0}\) is progressively measurable with respect to \((\mathcal{F}_t)_{t \geq 0}\) and satisfies the
Markov property

\[ E(f(X_t)|\mathcal{F}_u) = T_{t-u}f(X_u), \quad f \in L_2(\mu), \quad 0 \leq u < t. \]  

Let \( L \) be the generator of \( (T_t)_{t \geq 0} \) and \( \mathcal{D}(L) \) its domain of definition on \( L_2(\mu) \). Denote by \( L^0_2(\mu) \) the subspace of functions in \( L_2(\mu) \) with \( \int_X f \, d\mu = 0 \). For \( f \in L^0_2(\mu) \) and \( t \geq 0 \) let

\[ S_t(f) = \int_0^t f(X_s) \, ds. \]

For a more detailed description of this setting see e.g. [1].

**Definition 1.2.** We say that there is a martingale approximation to \( S_t(f) \) if there exist two processes \((M_t)_{t \geq 0}\) and \((A_t)_{t \geq 0}\) on \((\Omega, \mathcal{A}, P)\) such that

1. \( S_t(f) = M_t + A_t, \quad t \geq 0, \)
2. \((M_t)_{t \geq 0}\) is a square-integrable martingale with respect to \((\mathcal{F}_t)_{t \geq 0}\) with stationary increments and \( M_0 = 0, \)
3. \( E(A_t)^2/t \to 0 \) as \( t \to \infty. \)

Again note that once the filtration \((\mathcal{F}_t)_{t \geq 0}\) is fixed, a martingale approximation is uniquely determined a.s. As in the discrete-time case, using a CLT for martingales with stationary increments (a careful discussion of which can be found in [6]) the existence of a martingale approximation implies that

\[ \frac{S_t(f)}{\sqrt{t}} \xrightarrow{t \to \infty} N(0, \sigma^2(f)), \]

where

\[ \sigma^2(f) = EM_1^2 = \lim_{t \to \infty} ES_t(f)^2/t. \]

The problem of the validity of the CLT for general continuous-time Markov processes has been studied less intensively than for discrete-time chains, and there seem to be few results via martingale approximation. Bhattacharya [1] proved the continuous-time analogue of Gordin and Lifšić’s [13] result. He assumed that there exists a solution to Poisson’s equation

\[ f = -Lg, \quad g \in \mathcal{D}(L). \]

Write

\[ S_t(f) = g(X_t) - g(X_0) - \int_0^t Lg(X_s) \, ds + g(X_0) - g(X_t). \]

Using Dynkin’s formula

\[ T_t g - g = \int_0^t T_{t-s} L g \, ds, \quad g \in \mathcal{D}(L) \]

it can be shown that \((M_t)_{t \geq 0}, M_t = g(X_t) - g(X_0) - \int_0^t Lg(X_s) \, ds, \) is a martingale with stationary increments with respect to \((\mathcal{F}_t)\). Furthermore, we evidently have
For the asymptotic variance, Bhattacharya [1] gave the formula
\[ \sigma^2(f) = 2(f, g) \] where \( f = -Lg \).

Kipnis and Varadhan [17] extended this approach in the context of reversible processes by solving (2) approximately via the resolvent. Using their method, de Masi et al. [8] studied reversible processes under weaker integrability conditions on the function \( f \), and Olla [19] considered non-reversible processes via the symmetrized operator \((L + L^*)/2\). Here our main goal is to show that the method of Kipnis and Varadhan [17] is universal in a certain sense. In Section 2 we introduce the resolvent representation of \( S_t(f) \) and show that there exists a martingale approximation to \( S_t(f) \) if and only if the resolvent representation converges (for the definitions see Section 2). Corresponding results are also formulated for Markov chains. In Section 3 this is applied to prove the CLT for stationary Markov processes under a condition analogous to that used by Maxwell and Woodroofe [18] in the discrete-time setting. As an example we consider moving average processes in continuous time. Furthermore, we give a sufficient condition for the existence of a martingale approximation if the generator \( L \) is a normal operator on \( L^2_c(\mu) \).

2. Martingale approximation and the resolvent

Let us start this section by recalling the resolvent representation of \( S_t(f) \), as introduced in [17]. Given \( \varepsilon > 0 \) let
\[ R_\varepsilon f = (\varepsilon I - L)^{-1}f = \int_0^\infty e^{-\varepsilon t}T_t f \, dt, \quad f \in L^2(\mu), \]
be the resolvent. Since \( R_\varepsilon f \in \mathcal{D}(L) \), given \( f \in L^2_c(\mu) \) we let \( g_\varepsilon = R_\varepsilon f \) and decompose
\[ S_t(f) = M_{t,\varepsilon} + \varepsilon S_t(g_\varepsilon) + A_{t,\varepsilon}, \quad (4) \]
where
\[ M_{t,\varepsilon} = g_\varepsilon(X_t) - g_\varepsilon(X_0) - \int_0^t (Lg_\varepsilon)(X_s) \, ds, \quad A_{t,\varepsilon} = -g_\varepsilon(X_t) + g_\varepsilon(X_0). \]

For each \( \varepsilon > 0 \) the process \( (M_{t,\varepsilon})_{t \geq 0} \) is a square-integrable martingale with stationary increments and \( M_{0,\varepsilon} = 0 \).

**Definition 2.1.** The decomposition (4) of \( S_t(f) \) is called the *resolvent representation*. The resolvent representation is said to *converge* if

1. \( \varepsilon \| g_\varepsilon \|^2 \to 0, \varepsilon \to 0 \).
2. There exists a decreasing sequence \( \varepsilon_n \to 0 \) with \( \varepsilon_{n+1} \geq c \varepsilon_n \) for some \( c > 0 \) such that for each \( t \geq 0 \), \( M_{t,\varepsilon_n} \) converges as \( n \to \infty \) to a limit in \( L^2(\Omega, \mathcal{F}, P) \).

Although this definition is rather technical, its significance becomes clear in the following theorem.
Theorem 2.2. Let \((X_t)_{t \geq 0}\) be a progressively measurable stationary ergodic Markov process, defined on a probability space \((\Omega, \mathcal{A}, P)\), with state space \((X, \mathcal{B})\), strongly continuous contraction semigroup \((T_t)_{t \geq 0}\) and stationary initial distribution \(\mu\). Let \(f \in L^2(\mu)\) and \(S_t(f) = \int_0^t f(X_s) \, ds\). Then there exists a martingale approximation to \(S_t(f)\) if and only if the resolvent representation of \(S_t(f)\) converges. In either case the limit variance is given by

\[
\sigma^2(f) = \lim_{n \to \infty} 2n(g_{1/n} - T_{1/n}g_{1/n}, g_{1/n}).
\]

A similar result also holds for discrete-time Markov chains. For \(\varepsilon > 0\) set \(g_\varepsilon = ((1 + \varepsilon)I - Q)^{-1}f\), where \(I\) denotes the identity, so that \((1 + \varepsilon)g_\varepsilon - Qg_\varepsilon = f\). Then we obtain a decomposition

\[
S_n(f) = M_{n, \varepsilon} + \varepsilon S_n(g_\varepsilon) + A_{n, \varepsilon},
\]

where

\[
M_{n, \varepsilon} = \sum_{k=1}^n (g_\varepsilon(X_k) - (Qg_\varepsilon)(X_{k-1})), \quad A_{n, \varepsilon} = (Qg_\varepsilon)(X_0) - (Qg_\varepsilon)(X_n).
\]

Again (6) is called the resolvent representation, and its convergence is defined as in the continuous-time case (just replace \(t\) by \(n\) in Definition 2.1). The theorem now goes

Theorem 2.3. Let \((X_n)_{n \geq 0}\) be a stationary ergodic Markov chain, defined on a probability space \((\Omega, \mathcal{A}, P)\), with state space \((X, \mathcal{B})\), transition operator \(Q\) and stationary initial distribution \(\mu\). Let \(f \in L^2(\mu)\) and \(S_n(f) = \sum_{k=1}^n f(X_k)\). Then there exists a martingale approximation to \(S_n(f)\) if and only if the resolvent representation of \(S_n(f)\) converges. In either case the limit variance is given by

\[
\sigma^2(f) = 2 \lim_{\varepsilon \to 0} \langle g_\varepsilon, f \rangle - \|f\|^2.
\]

Remark 2.4. Wu and Woodroofe [21] studied approximations by triangular arrays \((M_{n,k})_{k \geq 1}\) of martingales (with respect to \(\mathcal{F}_k\)). If

\[
\max_{k \leq n} E(S_k - M_{n,k})^2 = o(\sigma_n^2),
\]

where \(\sigma_n^2 = ES_n^2(f) \to \infty\), \((M_{n,k})_{k \geq 1}\) is called a martingale approximation scheme. It is called stationary if for each \(n\), \((M_{n,k})_{k \geq 1}\) has stationary differences, and non-triangular if \(M_{n,k} = M_k\) does not depend on \(n\). In this terminology, the martingale approximations of Definition 1.1 are stationary and non-triangular martingale approximation schemes. Wu and Woodroofe [21, Theorem 1], obtained necessary and sufficient conditions for the existence of such martingale approximation schemes. However, the martingale approximation schemes they construct are either stationary or non-triangular but the proof of Theorem 1 in [21] does not yield the existence of martingale approximation schemes which are both stationary and non-
triangular. Moreover, the martingale approximation schemes constructed do not imply the CLT. In their paper they also showed that the validity of a conditional version of the CLT is equivalent to a Lindeberg-type condition for the martingale approximation scheme.

Let us turn to the proof of Theorem 2.2. We will first prove two lemmas. Let

\[ V_t f = \int_0^t T_s f \, ds, \quad f \in L_2^0(\mu). \]

**Lemma 2.5.** Suppose that \( |V_n f| = o(\sqrt{n}) \). Then \( \sqrt{\varepsilon} g \|g\| \rightarrow 0, \quad \varepsilon \rightarrow 0. \)

**Proof.** Observe that \( |V_t f| \leq |V_{[t]} f| + |f| \), where \([t]\) denotes the integer part of \(t\). Hence there is a non-increasing sequence \( \phi_n \rightarrow 0 \) such that \( |V_t f| / \sqrt{t} \leq \phi_{[t]}, \quad t \geq 1 \). Therefore it is easy to find a bounded, continuously differentiable function \( \psi \) on \([0, \infty)\) such that \( \psi(t) \geq \phi_{[t]}, \quad t \geq 1 \), and \( \psi(t) \rightarrow 0, \quad t \rightarrow \infty \). Using the formula

\[ g = \int_0^\infty \varepsilon e^{-\varepsilon t} V_t f \, dt \]  

for the resolvent we estimate

\[ \sqrt{\varepsilon} \|g\| \leq \int_0^1 \varepsilon^{3/2} e^{-\varepsilon t} \|V_t f\| \, dt + \int_1^\infty \varepsilon^{3/2} e^{-\varepsilon t} \|V_t f\| \, dt \]

\[ \leq \varepsilon^{3/2} \|f\| + \int_0^\infty \varepsilon^{3/2} e^{-\varepsilon t} \sqrt{t} \psi(t) \, dt. \]

Substituting \( u = \varepsilon t \) in the second term, we obtain \( \int_0^\infty e^{-u} \sqrt{u} \psi(u/\varepsilon) \, du \), which tends to 0 as \( \varepsilon \rightarrow 0 \) by dominated convergence. \( \Box \)

**Lemma 2.6.** For \( \varepsilon, \delta > 0 \) we have

\[ |\langle g_\varepsilon - g_\delta - T_t(g_\varepsilon - g_\delta), g_\varepsilon - g_\delta \rangle| \leq 2t(\varepsilon + \delta) (\|g_\varepsilon\|^2 + \|g_\delta\|^2). \]  

**Proof.** From Dynkin’s formula (3)

\[ g_\varepsilon - g_\delta - T_t(g_\varepsilon - g_\delta) = \int_0^t (T_s L g_\delta - T_s L g_\varepsilon) \, ds. \]

Hence

\[ \langle g_\varepsilon - g_\delta - T_t(g_\varepsilon - g_\delta), g_\varepsilon - g_\delta \rangle = \int_0^t \langle T_s L g_\delta - T_s L g_\varepsilon, g_\varepsilon - g_\delta \rangle \, ds. \]  

Now since \( \delta g_\delta - L g_\delta = f \)

\[ T_s L g_\delta = \delta T_s g_\delta - T_s f. \]
Therefore

\[ |(T_s Lg_\delta - T_s Lg_\varepsilon - g_\varepsilon - g_\delta)| = |(\delta T_s g_\delta - \varepsilon T_s g_\varepsilon, g_\varepsilon - g_\delta)| \]

\[ \leq \delta |(T_s g_\delta, g_\delta)| + \varepsilon |(T_s g_\varepsilon, g_\varepsilon)| \]

\[ + \delta |(T_s g_\delta, g_\varepsilon)| + \varepsilon |(T_s g_\varepsilon, g_\delta)| \]

\[ \leq \delta \|g_\delta\|^2 + \varepsilon \|g_\varepsilon\|^2 + (\varepsilon + \delta)\|g_\delta\|\|g_\varepsilon\| \]

\[ \leq 2(\varepsilon + \delta)(\|g_\varepsilon\|^2 + \|g_\delta\|^2). \]

Applying this inequality in (10) yields the result. \( \Box \)

**Proof of Theorem 1.** First assume that there exists a martingale approximation 
\( S_t(f) = M_t + A_t. \) Since \( M_0 = 0 \)

\[ \frac{1}{n} \|V_n f\|^2 = \frac{1}{n} E(E(S_n(f)|F_0))^2 \]

\[ = \frac{1}{n} E(E(A_n) | F_0))^2 \]

\[ \leq \frac{1}{n} EA_n^2 \rightarrow 0, \quad n \rightarrow \infty. \]

Thus \( \|V_n f\| = o(\sqrt{n}), \) and Lemma 2.5 applies. For any \( h \in L_2(\mu), \) from the Schwarz inequality

\[ E \left( \int_0^t h(X_s) \, ds \right)^2 \leq E \left( t \int_0^t h(X_s)^2 \, ds \right) = t^2 \|h\|^2. \]  

(11)

From (11) and Lemma 2.5 it follows that \( \varepsilon E(S_t(g_\varepsilon))^2 \rightarrow 0, \varepsilon \rightarrow 0, \) for any \( t > 0. \) Let us show that \( M_{t, \varepsilon} \) converges in \( L_2(\Omega, \mathcal{A}, P) \) along the sequence \( \varepsilon_n = 1/n \) to \( M_t. \) Since both \( (M_t)_{t \geq 0} \) and \( (M_{t, \varepsilon})_{t \geq 0} \) are martingales with stationary increments with respect to \( (F_t) \), so is \( (M_t - M_{t, \varepsilon}). \) Therefore

\[ E(M_{t, \varepsilon} - M_t)^2 = 1/n E(M_{m, \varepsilon} - M_m)^2 \]

\[ \leq 3/n EA_n^2 + 3/n EA_{m,n}^2 + 3/n \varepsilon^2 E(S_m(g_\varepsilon))^2 \]

By assumption, \( 1/n EA_n^2 \rightarrow 0 \) as \( n \rightarrow \infty. \) Furthermore, \( EA_{m,1/n}^2 \leq 4\|g_1/n\|^2, \) thus using Lemma 2.5, \( 1/n EA_{m,1/n}^2 \rightarrow 0. \) Finally from (11), \( E(S_n(g_{1/n})^2 \leq t^2 n^2 \|g_{1/n}\|^2, \) and we obtain the conclusion for the last term. This shows that the resolvent representation converges.

Conversely, assume that the resolvent representation converges. Since \( E(S_t(g_\varepsilon)^2) \leq t^2 \|g_\varepsilon\|^2, \) \( \varepsilon S_t(g_\varepsilon) \rightarrow 0 \) as \( \varepsilon \rightarrow 0 \) in \( L_2(\Omega, \mathcal{A}, P). \) From the resolvent representation (4),

\[ A_{t, \varepsilon} = S_t(f) - M_{t, \varepsilon} - \varepsilon S_t(g_\varepsilon). \]

Since for \( t > 0 \) both \( M_{t, \varepsilon} \) and \( \varepsilon_n S_n(g_{\varepsilon_n}) \) converge as \( n \rightarrow \infty \) in \( L_2(\Omega, \mathcal{A}, P), \) and \( S_t(f) \) does not depend on \( n, \) it follows that \( A_{t, \varepsilon} \) also converges in \( L_2(\Omega, \mathcal{A}, P). \) Let us show that in fact \( A_{t, \delta_n} \) converges along an arbitrary sequence \( \delta_k \rightarrow 0. \) Let \( n(k) \) be such that
\( \varepsilon_{n(k)+1} \leq \delta_k \leq \varepsilon_{n(k)} \). Then \( \delta_k \geq c \varepsilon_{n(k)} \). From (1) and Lemma 2.6,

\[
E(A_t, \delta_k - A_{t,n(k)})^2 = 2(g_{\delta_k} - g_{\varepsilon_{n(k)}} - T_i(g_{\delta_k} - g_{\varepsilon_{n(k)}})) \leq 4t(\delta_k + \varepsilon_{n(k)})(\|g_{\delta_k}\|^2 + \|g_{\varepsilon_{n(k)}}\|^2) \leq 8t \varepsilon_{n(k)} \|g_{\varepsilon_{n(k)}}\|^2 + 4t(1 + 1/c) \delta_k \|g_{\delta_k}\|^2 \to 0, \quad k \to \infty.
\]

(12)

Arguing with the resolvent representation as above it follows that both

\[
M_{t, \varepsilon} \to M_t \quad \text{and} \quad A_{t, \varepsilon} \to A_t, \quad \varepsilon \to 0 \quad \text{in} \quad L_2(\Omega, \mathcal{F}, P), \quad t \geq 0,
\]

where \((M_{t, \varepsilon})_{t \geq 0}\) is a martingale with stationary increments with respect to \( (\mathcal{F}_t) \), \( M_0 = 0 \) and \( EM_t^2 < \infty, EA_t^2 < \infty \) for every \( t \). Thus it remains to show that \( EA_t^2 / t \to 0 \). But

\[
EA_t^2 \leq 3EA_{t, \varepsilon}^2 + 3E(M_{t, \varepsilon} - M_t)^2 + 3\varepsilon^2 ES_t(f_{\varepsilon})^2.
\]

Now let \( \varepsilon = 1/t \) and proceed as in [17] in the discrete-time situation to obtain the conclusion. Therefore we have a martingale approximation to \( S_t(f') \).

Finally let us prove the formula for the limit variance. We have that

\[
\sigma^2(f) = EM_1^2 = \lim_{n \to \infty} EM_{1, 1/n}^2.
\]

Since \((M_{t, 1/n}, t \geq 0)\) is a martingale with stationary increments,

\[
EM_{1, 1/n}^2 = nEM_{1, 1/n}^2 = nE \left( g_{1/n}(X_{1/n}) - g_{1/n}(X_0) - \int_0^{1/n} (Lg_{1/n})(X_s) \, ds \right)^2.
\]

For any \( \varepsilon > 0 \), \( Lg_\varepsilon = -f + \varepsilon g_\varepsilon \), hence \( \|Lg_\varepsilon\| \leq \|f\| + \varepsilon \|g_{\varepsilon}\| \leq 2\|f\| \). Thus

\[
\frac{nE \left( \int_0^{1/n} (Lg_{1/n})(X_s) \, ds \right)^2}{1/n\|Lg_{1/n}\|^2} \to 0.
\]

Since \( nE(g_{1/n}(X_{1/n}) - g_{1/n}(X_0))^2 = 2n(g_{1/n} - T_{1/n}g_{1/n}, g_{1/n}) \), the formula for \( \sigma^2(f) \) follows. The theorem is thus proved. \( \square \)

**Proof of Theorem 2.** The proof is similar to that of Theorem 1, therefore we only show how to obtain the formula for the variance.

\[
\sigma^2(f) = \lim_{\varepsilon \to 0} EM_{1, \varepsilon}^2 = \lim_{\varepsilon \to 0} (\|g_{\varepsilon}\|^2 - \|Qg_{\varepsilon}\|^2).
\]

Furthermore,

\[
\|g_{\varepsilon}\|^2 - \|Qg_{\varepsilon}\|^2 = -2\varepsilon\|g_{\varepsilon}\|^2 - \varepsilon^2\|g_{\varepsilon}\|^2 + 2(g_{\varepsilon}, f) - \|f\|^2 + 2\varepsilon(g_{\varepsilon}, f).
\]

All terms vanish as \( \varepsilon \to 0 \) except for \( 2(g_{\varepsilon}, f) - \|f\|^2 \), and the formula for \( \sigma^2(f) \) follows. \( \square \)
3. Asymptotic normality

In the following theorem we apply Theorem 1 to prove the CLT for stationary Markov processes under a condition which is analogous to that used by Maxwell and Woodroofe [18] in a discrete-time setting.

**Theorem 3.1.** Let \((X_t)_{t \geq 0}\) be a progressively measurable stationary ergodic Markov process with state space \((X, \mathcal{B})\), strongly continuous contraction semigroup \((T_t)_{t \geq 0}\) and stationary initial distribution \(\mu\). Let \(f \in L^2_2(\mu)\) and \(S_t(f) = \int_0^t f(X_s) \, ds\). Suppose that \(f\) satisfies
\[
\int_1^\infty \|V_t(f)\|/t^{3/2} \, dt < \infty. \tag{13}
\]
Then there exists a martingale approximation to \(S_t(f)\). In particular \(S_t(f)/\sqrt{t}\) is asymptotically normal with variance \(\sigma^2(f)\) given in (5), and
\[
\sigma^2(f) = \lim_{t \to \infty} E S_t(f)^2 / t.
\]

**Proof.** We show that the resolvent approximation converges. Let \(\varepsilon_n = 1/2^n\). The main point is to show that
\[
\sum_{n \geq 1} \sqrt{\varepsilon_n} \sup_{\varepsilon_n \leq \varepsilon < \varepsilon_{n-1}} \|g_\varepsilon\| < \infty. \tag{14}
\]
From this it follows immediately that \(\varepsilon \|g_\varepsilon\|^2 \to 0, \varepsilon \to 0\). Furthermore, from (12)
\[
E(A_{t,\varepsilon_{n+1}} - A_{t,\varepsilon_n})^2 \leq 8t\varepsilon_n \|g_{\varepsilon_n}\|^2 + 12t\varepsilon_{n+1} \|g_{\varepsilon_{n+1}}\|^2,
\]
and since \(\sqrt{a+b} \leq \sqrt{a} - \sqrt{b}, a, b \geq 0\),
\[
\|A_{t,\varepsilon_{n+1}} - A_{t,\varepsilon_n}\|_{L_2(\Omega, \mathcal{F}, P)} \leq C_1\sqrt{\varepsilon_n} \|g_{\varepsilon_n}\| + C_2 \sqrt{\varepsilon_{n+1}} \|g_{\varepsilon_{n+1}}\|.
\]
Therefore, from (14)
\[
\sum_{n \geq 1} \|A_{t,\varepsilon_{n+1}} - A_{t,\varepsilon_n}\|_{L_2(\Omega, \mathcal{F}, P)} < \infty,
\]
and \(A_{t,\varepsilon_n}\) converges in \(L_2(\Omega, \mathcal{F}, P)\) as \(n \to \infty\). From the resolvent representation it follows that \(M_{t,\varepsilon_n}\) also converges. It remains to show (14). Given \(\varepsilon > 0\) choose \(n\) such that \(\varepsilon_n \leq \varepsilon < \varepsilon_{n-1}\). From (8)
\[
\|g_\varepsilon\| \leq \varepsilon \int_0^1 e^{-\varepsilon t} \|V_t(f)\| \, dt + \varepsilon \int_1^\infty e^{-\varepsilon t} \|V_t(f)\| \, dt
\]
\[
\leq 2\varepsilon_n \|f\| + 2\varepsilon_n \int_1^\infty e^{-\varepsilon_n t} \|V_t(f)\| \, dt.
\]
Hence
\[
\sum_{n \geq 1} \sqrt{\varepsilon_n} \sup_{\varepsilon_n \leq \varepsilon < \varepsilon_{n-1}} \|g_\varepsilon\| \leq 2 \sum_{n \geq 1} \varepsilon_n^{3/2} \|f\| + 2 \int_1^\infty \|V_t(f)\| \left( \sum_{n \geq 1} \varepsilon_n^{3/2} e^{-\varepsilon_n t} \right) \, dt.
\]
But \( \sum_{n \geq 1} \frac{e^{3/2}}{n^3} e^{-ui} \) can be seen to be \( O(t^{-3/2}) \) (cf. [18, p. 715]), and condition (13) implies (14).  

**Corollary 3.2.** If \( f \in L^0_2(\mu) \) satisfies

\[
\int_1^\infty \frac{\|T_t(f)\|}{\sqrt{t}} \, dt < \infty,
\]

then it also fulfills (13), and therefore the conclusions of the theorem remain valid.

**Example 3.3** (Moving average processes). We consider the semigroup of translation operators on \( L^2[0, \infty) \) given by

\[
(T_t f)(u) = f(u + t), \quad f \in L^2[0, \infty)
\]

and denote the generator of \((T_t)\) by \( L \). Let \((\xi_t)_{t \in \mathbb{R}}\) be a square-integrable, real-valued process with stationary, independent increments, \( E\xi_t = 0 \) and \( E\xi_t^2 = dt \) (cf. [7, p. 111]). Each \( f \in L^2[0, \infty) \) gives rise to a stationary, ergodic process \((Y_t)_{t \in \mathbb{R}}\), defined by the stochastic integrals

\[
Y_t(f) = Y_t = \int_{-\infty}^t f(t - s) \, d\xi_s.
\]

Let \( \mathcal{F}_s = \sigma(\xi_u, u \leq s) \). We have that

\[
E(Y_t | \mathcal{F}_u) = \int_{-\infty}^u f(t - s) \, d\xi_s = \int_{-\infty}^u T_{t-u} f(u - s) \, d\xi_s, \quad t \geq u,
\]

hence

\[
E(E(Y_t | \mathcal{F}_0))^2 = \|T_t f\|^2 = \int_t^\infty f(u)^2 \, du.
\]

Although \((Y_t)_{t \geq 0}\) is not constructed from a Markov process in the way discussed above, these considerations show that our method can still be used with the translation semigroup \((T_t)\) in place of the semigroup of the Markov process. For example, the martingales \((M_{t,\varepsilon})_{t \geq 0}, \varepsilon > 0\), now take the form

\[
M_{t,\varepsilon} = Y_t(g_{\varepsilon}) - Y_0(g_{\varepsilon}) - \int_0^t Y_s(Lg_{\varepsilon}) \, ds,
\]

where \( g_{\varepsilon} \) is formed via the semigroup \((T_t)\). Thus Corollary 3.2 applies, and we obtain that if \( f \in L^2[0, \infty) \) satisfies

\[
\int_1^\infty \frac{1}{\sqrt{t}} \int_t^\infty f(u)^2 \, du \, dt < \infty,
\]

then

\[
\frac{1}{\sqrt{t}} \int_0^t Y_s \, ds \xrightarrow{t \to \infty} N(0, \sigma^2(f)),
\]

where \( \sigma^2(f) \) is given by (5).
Remark 3.4. Finally suppose that the generator $L$ of the Markov process is a normal operator on $L^2_\mu(\mathbb{R})$. Although conditions for the CLT in the discrete-time case, i.e. in case of a normal transition operator $Q$, have been studied intensively (cf. Gordin and Lifšic [14], who announced the result and later published the complete proofs in [5, Section IV.7], or Derriennic and Lin [9,10]), there seems to be no continuous-time version in the literature so far. Let us formulate a sufficient condition for the convergence of the resolvent representation and hence for the validity of the CLT for such operators. Given $f \in L^2_\mu(\mathbb{R})$ let $\rho_f$ be the spectral measure of $L$ with respect to $f$ (cf. [3, pp. 123–125]), and let $\sigma(L)$ denote the spectrum of $L$. One can show that if

$$\int_{\sigma(L)} \frac{1}{|z|^2} \rho_f(\mathrm{d}z) < \infty, \quad (15)$$

then the resolvent representation converges and hence there exists a martingale approximation to $S_t(f)$, where the limit variance is given by

$$\sigma^2(f) = -2 \int_{\sigma(L)} \frac{1}{z} \rho_f(\mathrm{d}z).$$

A rich class of examples of Markov processes with normal but not necessarily self-adjoint generator arises from convolution semigroups on compact, commutative hypergroups (cf. [4]). Further details can be found in [15]. As pointed out in [10] for the discrete-time case, the condition (15) is weaker than (13) as used in Theorem 3.1.

Acknowledgements

This paper is based on a section of my Ph.D. thesis (Universität Göttingen, 2004), and I would like to thank my supervisor Professor M. Denker for constant encouragement, discussions and advice. Further I thank the referee for pointing out several references.

References