Averaging for Markov Chains: The Convergence Theorem

16.1 Introduction

In this chapter we prove a result concerning averaging for Markov chains. The techniques presented lead to a weak-convergence-type result showing that expectations under the original chain and under the averaged chain are close. The technique is to work with the backward equation for the two Markov chains. The fundamental estimate (5.2.2) plays a central role. This estimate is analogous to the maximum principle for parabolic PDEs. In Chapter 20 we use techniques similar to those in this chapter, based on the maximum principle, to prove a homogenization result for parabolic PDEs. The main theorem is stated in Section 16.2 and is proved in Section 16.3. The chapter concludes with bibliographical notes in Section 16.4.

16.2 The Theorem

The setup is as in Chapter 9. To make the proofs transparent we concentrate on the finite state space case. Let \( \mathcal{I}_x, \mathcal{I}_y \subseteq \{1, 2, \cdots\} \) be finite sets. Consider a continuous-time Markov chain \( z(t) = (x(t), y(t)) \) on \( \mathcal{I}_x \times \mathcal{I}_y \). We assume that this Markov chain is parameterized by \( \varepsilon \) and that the backward equation has the form

\[
\frac{dv}{dt} = \frac{1}{\varepsilon} Q_0 v + Q_1 v \quad (16.2.1)
\]

where \( Q_0, Q_1 \) are given by (9.2.4). Let \( X(t) \) be a Markov chain on \( \mathcal{I}_x \) with backward equation

\[
\frac{dv_0}{dt} = Q_1 v_0, \quad (16.2.2)
\]
and with $\bar{Q}_1$ given by (9.3.1). We are interested in approximating $x(t)$ by $X(t)$. Note that the formula for the approximate process implied by the Kolmogorov equation is exactly that derived in Chapter 9 by means of formal asymptotics.

Note that $x(t)$ is not itself Markovian; only the pair $(x(t), y(t))$ is. Thus we are approximating a non-Markovian stochastic process by a Markovian one. To be precise, we prove that, at any fixed time, the statistics of $x(t)$ are close to those of $X(t)$. That is, we prove weak convergence of $x(t)$ to $X(t)$ at any fixed time $t$.

**Theorem 16.1.** For any $t > 0$, $x(t) \Rightarrow X(t)$, as $\varepsilon \to 0$.

### 16.3 The Proof

Let $v_0$ be defined as in (16.2.2). We then have

$$v_0 \in \mathcal{N}(Q_0), \quad \frac{dv_0}{dt} - Q_1 v_0 \perp \mathcal{N}(Q_0^T),$$

by construction. Hence there exists $v_1$ so that

$$Q_0 v_0 = 0,$$

$$Q_0 v_1 = \frac{dv_0}{dt} - Q_1 v_0.$$

We can make $v_1$ unique by insisting that it is orthogonal to the null space of $Q_0^*$, although this particular choice is not necessary. We simply ask that a solution is chosen that is bounded, with bounded derivative in time.

For any such $v_1$ and for $v_0$ given by (16.2.2), define

$$r = v - v_0 - \varepsilon v_1.$$

Substituting $v = v_0 + \varepsilon v_1 + r$ into (16.2.1) and using the properties of $v_0, v$, we obtain

$$\frac{dv_0}{dt} + \varepsilon \frac{dv_1}{dt} + \frac{dr}{dt} = \frac{1}{\varepsilon} Q_0 v_0 + Q_0 v_1 + \frac{1}{\varepsilon} Q_0 r + Q_1 v_0 + \varepsilon Q_1 v_1 + Q_1 r.$$

Hence

$$\frac{dr}{dt} = \left(\frac{1}{\varepsilon} Q_0 + Q_1\right) r + \varepsilon q,$$

$$q = Q_1 v_1 - \frac{dv_1}{dt}.$$

Now $Q := \frac{1}{\varepsilon} Q_0 + Q_1$ is the generator of a Markov chain. Hence using $| \cdot |_{\infty}$ to denote the supremum norm on vectors over the finite set $\mathcal{I}_x \times \mathcal{I}_y$, as well as the induced operator norm, we have

$$|e^{Qt}|_{\infty} = 1.$$  \hspace{1cm} (16.3.1)