Rates of Convergence in the Central Limit Theorem for Markov Chains

Marc Corluy

July 31, 2008

Motivation

Most non-deterministic phenomena are modeled by continuous processes, often Brownian motion related.

Some phenomena do not allow modeling with continuous processes.

- Financial Mathematics : Occurrence of War, Pestilence et al.
- Quantum Mechanics : Discontinuity is inherent in its basic premises.

Solution : Use Processes with Jumps in your modeling.

Caveat: In order to have good modeling, you might be forced to use processes where the magnitude of the jumps is not bounded.

The Markov Chain

Framework

1. Let Y_n be a symmetric Markov chain on $\mathbb{Z}^d$.
2. Y_n can have unbounded range. i.e. for every K, there exists x and y s.t. $\|x - y\| > K$ and $\mathbb{P}(Y_{n+1} = y \mid Y_n = x) > 0$.
3. C_{xy} be the conductances between x and y; the transition probabilities for the Markov Chain Y_n are defined by

$$p(x, y) = \mathbb{P}^x(Y_1 = y) = \frac{C_{xy}}{\sum_z C_{xz}} \tag{1.1}$$

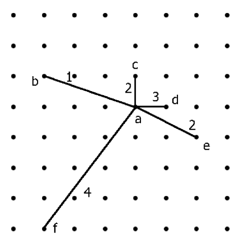
where every $C_{xy} \in [0, \infty)$.

Example (Space Homogeneous Case)

$$\mathbb{P}(Y_{n+1} = y \mid Y_n = x) = \frac{c}{\|y - x\|^{d+\gamma}} \tag{1.2}$$

with $\gamma \in (0, 2)$ and $c > 0$ a normalization factor.

Example (Grid in $\mathbb{Z}^2$)



In this grid

$$\sum_{y \in \mathbb{Z}^2} C_{ay} = 1 + 2 + 3 + 2 + 4 = 12$$

and therefore

$p(a, b) = \frac{1}{12}$, $p(a, c) = \frac{2}{12}$, $p(a, d) = \frac{3}{12}$, $p(a, e) = \frac{2}{12}$, and $p(a, f) = \frac{4}{12}$.

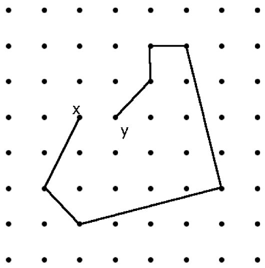
Conditions on the Conductances

(A0) $\forall x, y \in \mathbb{Z}^d : C_{xy} = C_{yx}$

(A1) $\forall x \in \mathbb{Z}^d, \exists c_1, c_2 > 0$ s.t. $c_1 \leq \sum_{y \in \mathbb{Z}^d} C_{xy} \leq c_2$

(A2) There exist $M_o \geq 1, \delta > 0$ s.t. for any x, y with $\|x - y\| = 1$, there exist $N > 2$ and $x_1, \ldots, x_N$ in $B(x, M_o)$ s.t. $x_1 = x$ and $x_N = y$ and $C_{x_i x_{i+1}} \geq \delta$ for $i = 1, \ldots, N-1$.

Example in $\mathbb{Z}^2$:



Conditions on the Conductances

(A3) There exists a decreasing function

$$\Phi : \mathbb{N} \to \mathbb{R}^+ \text{ with } \sum_{i=1}^{\infty} i^{d+1}\Phi(i) < \infty \qquad (1.3)$$

s.t. for all x, y : $C_{xy} \leq \Phi(\|x-y\|)$.

Note : (A3) states that the C_{xy} satisfy a uniform finite second moment condition:

$$\begin{aligned}
\sum_{y\in\mathbb{Z}^d} \|x-y\|^2 C_{xy} &\leq \sum_{y\in\mathbb{Z}^d} \|x-y\|^2 \Phi(\|x-y\|) \\
&\leq \sum_{i=0}^{\infty} \sum_{i<\|x-y\|\leq i+1} \|x-y\|^2 \Phi(\|x-y\|) \\
&\leq \sum_{i=0}^{\infty} (i+1)^2 \Phi(i) \sum_{i<\|x-y\|\leq i+1} 1 \\
&\leq c_3 \sum_{i=0}^{\infty} (i+1)^2 \Phi(i)(i+1)^{d-1} \\
&< \infty \qquad (1.4)
\end{aligned}$$

Conditions on the Conductances

(A4) Define $A^{(h)} : h\mathbb{Z}^d \to \mathbb{R}$ as

$$\left[A^{(h)}(x)\right]_{ij} = \frac{1}{2} \sum_{z \in \mathbb{Z}^d} z_i \; z_j \; C_{\frac{x}{h}, \frac{x}{h}+z}. \tag{1.5}$$

Note that $x \in h\mathbb{Z}^d$, so $C_{\frac{x}{h}, \frac{x}{h}+z}$ is a conductance on the grid $\mathbb{Z}^d$. Let there be a Borel measurable function $A : \mathbb{R}^d \to M_{d \times d}$ such that A is symmetric and uniformly elliptic where the map $x \to A(x)$ is continuous, and $A^{(h)}$ needs to converge to A such that for some $c > 0$ and $\beta > 0$,

$$\|A - A^{(h)}\|_\infty \leq ch^\beta. \tag{1.6}$$

(A5) Let C_{xy} be Hölder continuous with exponent α. More precisely, there exist $c > 0$ and $\alpha > 0$ such for all x and y,

$$|C_{xy_1} - C_{xy_2}| \leq c \, \|y_1 - y_2\|^\alpha. \tag{1.7}$$

Definition

The rescaled Markov Chain $X_t^{(h)}$ is defined as

$$X_t^{(h)} := hY_{\frac{t}{h^2}} \tag{1.8}$$

Note that the state space of this process is $h\mathbb{Z}^d$.

Notation

Let $\mathscr{C}([0, t_o], \mathbb{R}^d)$ denote the collection of all continuous paths from $[0, t_o]$ to $\mathbb{R}^d$. The space $\mathscr{D}([0, t_o], \mathbb{R}^d)$ is the collection of all cadlag functions, i.e. all paths that are continuous on the right and have left limits.

Theorem (Bass and Kumagai, 2008)

Suppose that (A0) to (A3) hold and let $A^{(h)}$ converge to A on compact sets. Also, put $[x]_h := h\left(\left[\frac{1}{h}x_1\right], \ldots, \left[\frac{1}{h}x_d\right]\right)$. Then the following holds:

1. *For each x and for each t_o the $\mathbb{P}^{[x]_h}$-law of $\{X_t^{(h)}\}_{0 \le t \le t_o}$ converges weakly with respect to the topology of the space $\mathscr{D}([0, t_o], \mathbb{R}^d)$. The limit gives full measure to $\mathscr{C}([0, t_o], \mathbb{R}^d)$.*
2. *If X_t is the canonical process on $\mathscr{C}([0, t_o], \mathbb{R}^d)$ and $\mathbb{P}^x$ is the weak limit of the $\mathbb{P}^{[x]_h}$-laws of $X_t^{(h)}$, then the process $\{X_t, \mathbb{P}^x\}$ has continuous paths and is the symmetric process corresponding to the Dirichlet form*

$$\mathcal{E}_A(f, f) = \int_{\mathbb{R}^d} \langle \nabla f(x) \mid A(x) \nabla f(x) \rangle \, dx \qquad (1.9)$$

Question

A natural question to ask is how fast this convergence occurs. In other words, estimate

$$|\mathbb{E}\phi(X_t^{(h)}) - \mathbb{E}\phi(X_t)| \leq F(h,t)\ \|\phi\| \tag{2.1}$$

for some function F.

Answer

In general, for any Markov process Y_t, there is a semigroup $\{\mathcal{P}_t\}_t$ such that $\mathcal{P}_t\phi(x) = \mathbb{E}\phi(Y_t^x)$ for any ϕ in the domain of $\mathcal{P}_t$. Thus, finding F for (2.1) transforms into finding F for

$$\|P_t\phi - P_t^h\phi\| \leq F(h,t)\ \|\phi\| \tag{2.2}$$

where $P_t\phi(x) = [e^{tL}\phi](x)$ and $P_t^h\phi(x) = [e^{tL^h}\phi](x)$ with

$$\begin{aligned} L\phi(x) &= \nabla \cdot A(x)\nabla\ \phi(x) \\ L^h\phi(x) &= \frac{1}{h^2}\sum_{z\in\mathbb{Z}^d}(\phi(x+hz)-\phi(x))\ C_{\frac{x}{h},\frac{x}{h}+z} \end{aligned} \tag{2.3}$$

Theorem (Duhamel's Formula)

Put $P_t = e^{tL}$ and $P_t^h = e^{tL^h}$. We can then perform the following calculation:

$$P_t\phi - P_t^h\phi = \int_0^t \frac{d}{ds}\left(P_t P_{t-s}^h \phi\right)\, ds \tag{2.4}$$

$$= \int_0^t \left(\frac{dP_s}{ds} \circ P_{t-s}^h\phi + P_s \circ \frac{dP_{t-s}^h}{ds}\phi\right)\, ds \tag{2.5}$$

$$= \int_0^t \left(LP_sP_{t-s}^h\phi - P_sL^hP_{t-s}^h\phi\right)\, ds \tag{2.6}$$

$$= \int_0^t P_s \circ (L - L^h) \circ P_{t-s}^h\phi \, ds \tag{2.7}$$

The end result is usually referred to as Duhamel's formula.

Problem

$L \circ P_{t-s}^h \phi$ will unavoidably involve terms of the form

$$\lightning \ \nabla P_t^h \phi(x) \ \lightning$$

Solution

We will have to introduce smooth versions of L and L^h. So,

$$\tilde{L}^h f(x) = \frac{1}{h^2} \sum_{z \in \mathbb{Z}^d} (f(x+hz) - f(x)) \, \tilde{C}^{\varepsilon}_{\frac{x}{h}, \frac{x}{h}+z} \tag{2.8}$$

where

$$\tilde{C}^{\varepsilon}_{x,y} = \int_{\mathbb{R}^d} \eta_\varepsilon(w) C_{x,y-w} \, dw, \tag{2.9}$$

and $L\phi(x) = \nabla \cdot \tilde{A}(x) \nabla \ \phi(x)$ where $\tilde{a}_{ij}(x) = \int_{\mathbb{R}^d} \eta_\varepsilon(w) a_{ij}(x) \, dw$.

The Price

We will have to find $F(h,t)$ for

$$\|P_t\phi - \tilde{P}_t\phi\| + \|\tilde{P}_t\phi - \tilde{P}_t^h\phi\| + \|\tilde{P}_t^h\phi - P_t^h\phi\| \;\leq\; F(h,t)\,\|\phi\|, \quad (2.10)$$

and each part has to be bound separately.

Theorem

If (A1) to (A4) are satisfied,

$$\|P_t^h - \tilde{P}_t^h\|_2 \;\leq\; c(t)\varepsilon^\alpha \quad (2.11)$$

Theorem (Chen, Qian, Hu, and Zheng, 1998)

If (A1) to (A4) are satisfied,

$$\|P_t - \tilde{P}_t\|_2 \;\leq\; c(t)\varepsilon^\alpha \quad (2.12)$$

Norms

We introduce the following norm:

$$\|\phi\|_{C_k} := \max_{1\le j\le k} \sup_{x\in\mathbb{R}^d} \left| \frac{\partial^j \phi}{\partial x_{i_1} \dots x_{i_j}}(x) \right|. \tag{2.13}$$

We also introduce

$$\|b'\|_\infty = \max_{1\le i,j=1\le d} \sup_{x\in\mathbb{R}^d} \left| \frac{\partial b^i}{\partial x_j} \right|, \tag{2.14}$$

and

$$\|b''\|_\infty = \max_{1\le i,j,k=1\le d} \sup_{x\in\mathbb{R}^d} \left| \frac{\partial^2 b^i}{\partial x_j \partial x_k} \right|, \tag{2.15}$$

as convenient shorthands, and we set

$$\|Y_t\|_\Delta := \max_{1\le j\le d} \sup_{x\in\mathbb{R}^d} \left| Y_t^{x,(j)}(\omega) \right|. \tag{2.16}$$

Remark (Flows)

In our estimates we will repeatedly encounter first and second order derivatives of functions of the form $f(x) = \mathbb{E}\phi(Y_t^x)$. A symbolic use of the chain rule would lead to

$$\frac{\partial}{\partial x}\mathbb{E}\ \phi(Y_t^x) \ = \ \mathbb{E}\ \phi'(Y_t^x)\ DY_t^x. \tag{2.17}$$

As it turns out, DY_t^x is a process that indeed behaves like a derivative in the sense that

$$Y_t^y - Y_t^x \ = \ \int_x^y DY_t^\xi\ d\xi \tag{2.18}$$

A rigorous treatment of so defined Derivatives for Ito Diffusions is widely available in the literature. We will revisit the topic for Integrators with Jumps later.

Theorem

If (A1) to (A5) are satisfied, then

$$\|\tilde{P}_t^h\phi - \tilde{P}_t\phi\|_\infty$$

$$\leq c_1\ \varepsilon^{-1}\ h^\beta\ \|\phi\|_{C_2} \sum_{i=1}^{d} \int_0^t \mathbb{E}\,\|\ D_i\tilde{X}_s\ \|_\Delta\ ds$$

$$+\ c_2\ h^\beta\ \|\phi\|_{C_1} \sum_{i,j=1}^{d} \int_0^t \mathbb{E}\,\|\ D_{ij}^2\tilde{X}_s\ \|_\Delta\ ds$$

$$+\ c_3\ h^\beta\ \|\phi\|_{C_2} \sum_{i,j=1}^{d} \int_0^t \max_{1\leq k,l\leq d} \sup_{x\in\mathbb{R}^d} \mathbb{E}\left|D_i\tilde{X}_s^{x,(k)} D_j\tilde{X}_s^{x,(l)}\right| ds$$

Theorem

If $\tilde{X}_t^x$ is a weak solution of

$$\tilde{X}_t^x = x + \int_0^t b(\tilde{X}_s^x)\ ds + \int_0^t \sigma(\tilde{X}_s^x) dW_s \tag{2.19}$$

then,

$$\mathbb{E}\|D_k\tilde{X}_t\|_\Delta \ \leq\ c_o\ e^{\left(c_1\|b'\|_\infty^2 + c_2\|\sigma'\|_\infty^2\right)\ t}, \tag{2.20}$$

and

$$\mathbb{E}\|D_{km}^2\tilde{X}_t\|_\Delta \ \leq\ \sqrt{c_{ob}\|b''\|_\infty^2 + c_{o\sigma}\|\sigma''\|_\infty^2}\ \sqrt{t}\ e^{\left(c_{1b}\|b'\|_\infty^2 + c_{1\sigma}\|\sigma'\|_\infty^2\right)\ t}. \tag{2.21}$$

Theorem

If (A1) to (A5) are satisfied, then

$$\|\tilde{P}_t^h\phi - \tilde{P}_t\phi\|_\infty \;\leq\; \left(\frac{c_1}{\varepsilon}t \;+\; \frac{c_2}{\varepsilon^2}\sqrt{\frac{c_{0b}}{\varepsilon^2}+1}\; t^{\frac{3}{2}} \;+\; c_3 t\right) \cdot \exp\left(\frac{c_4}{\varepsilon^2}\left(\frac{c_5}{\varepsilon^2}+1\right)t\right) \cdot\; h^\beta \;\|\phi\|_{C_2} \quad (2.22)$$

Lemma

If $\mathcal{P}_t$ and $\mathcal{Q}_t$ are Markov semigroups, then, for any ϕ that is in both domains,

$$\|\mathcal{P}_t\phi - \mathcal{Q}_t\phi\|_2 \;\leq\; \sqrt{2\,\|\,\mathcal{P}_t\phi - \mathcal{Q}_t\phi\,\|_\infty\,\|\phi\|_1}. \qquad (2.23)$$

Theorem

If (A1) to (A5) are satisfied, then

$$\|P_t\phi - P_t^h\phi\|_2 \leq c_1 \, h^{\frac{\alpha\beta}{2\alpha+2}} e^{\left(c_2 h^{4-\frac{2\beta}{\alpha+2}}\right)} \cdot \max\left\{\|\phi\|_2, \frac{\|\phi\|_{C_2} + \|\phi\|_1}{2}\right\} \tag{2.24}$$

Remark

Note that for $h < 1$, which is the only important part for convergence,

$$\exp\left(c_2 h^{4-\frac{2\beta}{\alpha+2}}\right) \leq c_2', \tag{2.25}$$

as $0 < \alpha, \beta < 1$. So for $h < 1$, we have

$$\|P_t\phi - P_t^h\phi\|_2 \leq c \, h^{\frac{\alpha\beta}{2\alpha+2}} \cdot \max\left\{\|\phi\|_2, \frac{\|\phi\|_{C_2} + \|\phi\|_1}{2}\right\}. \tag{2.26}$$

Jump Processes as Integrator

Set

$$X_t^x = x + \int_0^t \int_S \Gamma(X_{s-}^x, z)(\mu - \nu)(ds, dz) \tag{3.1}$$

where S is the state space, Γ a measurable function from $\mathbb{R} \times S$ to $\mathbb{R}$, μ is a Poisson point process, and $\nu(A) = t\lambda(A)$ where λ is the Lebesgue measure on $\mathbb{R}^d$. We can formally write

$$DX_t^x = 1 + \int_0^t \int_S \Gamma'(X_{s-}^x, z) DX_{s-}^x (\mu - \nu)(ds, dz) \tag{3.2}$$

where

$$X_t^x - X_t^y = \int_y^x DX_t^u du. \tag{3.3}$$

This calculation is made rigorous within the framework of flows.

Definition (Stochastic Flow)

The flow associated with a stochastic differential equation is the family of stochastic variables defined by

$$\Phi_X = \left\{ X_t^x \mid x \in \mathbb{R}^d, t \in [0, +\infty), X_t^x \text{ solves the SDE} \right\}. \quad (3.4)$$

Conditions on Γ

(G1) $\int_S |\Gamma(x,z) - \Gamma(y,z)|^{2p} \lambda(dz) \leq C|x-y|^{2p}$

(G2) $\int_S |\Gamma'(x,z)|^{2p} \lambda(dz) \leq C$

(G3) $\int_S |\Gamma'(x,z) - \Gamma'(y,z)|^{2p} \lambda(dz) \leq C|x-y|^{2p}$

(G4) $\int_S |\Gamma''(x,z)|^{2p} \lambda(dz) \leq C$

Theorem

If Γ satisfies (G1), then there exists a version of X_t^x that is jointly cadlag in t and continuous in x.

Theorem

If Γ satisfies (G1) and (G2), then the SDE

$$Y_t^x = 1 + \int_0^t \int_S \Gamma'(X_{s-}^x, z) Y_{s-}^x (\mu - \nu)(ds, dz) \tag{3.5}$$

has a unique solution DX_t^x which has moments of all orders and there exists a version of DX_t^x that is jointly cadlag in t and continuous in x.

Theorem

If Γ satisfies (G1) to (G4), then for all x and y,

$$X_t^x - X_t^y = \int_y^x DX_t^\xi d\xi. \tag{3.6}$$

Faà di Bruno's Formula

For all $n > 0$,

$$\frac{d^n}{dx^n} f(g(x)) = \sum_{k=1}^{n} f^{(k)}(g(x)) B_{n,k}\left(g'(x), g''(x), \dots, g^{(n-k+1)}(x)\right), \tag{3.7}$$

where the Bell polynomial is defined as follows:

$$B_{n,k}(x_1, x_2, \dots, x_{n-k+1}) = \sum \frac{n!}{j_1! j_2! \cdots j_{n-k+1}!} \left(\frac{x_1}{1!}\right)^{j_1} \left(\frac{x_2}{2!}\right)^{j_2} \cdots \left(\frac{x_{n-k+1}}{(n-k+1)!}\right)^{j_{n-k+1}} \tag{3.8}$$

The sum extends over all sequences $j_1, \dots, j_n$ of non-negative integers such that $\sum_{i=1}^{n} j_i = k$ and $\sum_{i=1}^{n} i j_i = n$.

Definition

We define the higher order derivatives $(n \geq 2)$ of X_t^x, as

$$D^n X_t^x = \int_0^t \int_S \sum_{k=1}^{n} \Gamma^{(k)}(X_{s-}^x, z) B_{n,k}(D^1 X_{s-}^x, \ldots, D^{n-k+1} X_{s-}^x)(\mu - \nu)(ds, dz)$$

Conditions on Γ

$(G1^{(n)})$ $\int_S |\Gamma^{(n)}(x,z) - \Gamma^{(n)}(y,z)|^{2p} \lambda(dz) \leq C|x-y|^{2p}$

$(G2^{(n)})$ $\int_S |\Gamma^{(n)}(x,z)|^{2p} \lambda(dz) \leq C$

Theorem

If Γ satisfies $(G1^{(1)})$ to $(G1^{(n)})$, and $(G2^{(1)})$ to $(G2^{(n)})$, then the SDE

$$Y_t^x = \int_0^t \int_S \Gamma'(X_{s-}^x, z) Y_{s-}^x \, (\mu - \nu)(ds, dz)$$

$$+ \int_0^t \int_S \sum_{k=2}^n \Gamma^{(k)}(X_{s-}^x, z) \; B_{n,k}(D^1 X_{s-}^x, \ldots, D^{n-k+1} X_{s-}^x) \; (\mu - \nu)(ds, dz)$$

has a unique solution $D^n X_t^x$ which has moments of all orders and there exists a version of $D^n X_t^x$ that is jointly cadlag in t and continuous in x.

Theorem

If Γ satisfies $(G1^{(1)})$ to $(G1^{(n+1)})$ and $(G2^{(1)})$ to $(G2^{(n+1)})$, then for all x and y,

$$D^{n-1}X_t^x - D^{n-1}X_t^y = \int_y^x D^n X_t^\xi d\xi. \tag{3.9}$$

Theorem (Taylor Expansion)

For any $n > 0$, we have the Taylor expansion

$$X_t^x = \sum_{k=0}^{n-1} \frac{(x - x_o)^k}{k!} D^k X_t^{x_o} \; + \; R_n(t; x, x_o), \tag{3.10}$$

where

$$R_n(t; x, x_o) \; = \; \int_{x_o}^{x}\int_{x_o}^{x_1} \cdots \int_{x_o}^{x_{n-1}} \; D^n X_t^{x_n} \; dx_n \dots dx_1 dx_o. \tag{3.11}$$

Furthermore, for almost every path ω, there exists an $\eta(\omega) \in [x_o \wedge x, x_o \vee x]$ such that

$$R_n(t; x, x_o) \; = \; D^n X_t^{\eta} \, \frac{(x - x_o)^n}{n!}. \tag{3.12}$$

Remark (Generalization)

Let X_t be defined as

$$X_t^x = x + \int_0^t b(X_{s-}^x)ds + \int_0^t \sigma(X_{s-}^x)dW_s + \int_0^t\int_S \Gamma(x_{s-}^x, z)(\mu - n)(ds, dz), \quad (3.13)$$

and let b, σ and Γ satisfy

$(\Phi 1^{(k)})$ $\int_S |\Phi^{(k)}(x,z) - \Phi^{(k)}(y,z)|^{2p}\lambda(dz) \leq C|x-y|^{2p}$

$(\Phi 2^{(k)})$ $\int_S |\Phi^{(k)}(x,z)|^{2p}\lambda(dz) \leq C$

for every $0 \leq k \leq N+1$ where N is the highest order derivative wanted, then for every $0 < l \leq N$, the previous theorems are also valid.

References

- Richard F. Bass, *Diffusions and Elliptic Operators*, Probability and its Applications, ISBN 0-387-98315-5, Springer Verlag, 1998
- Richard F. Bass and Takashi Kumagai, *Symmetric Markov Chains on $\mathbb{Z}^d$ with Unbounded Range*, Transactions of the American Mathematical Society, Vol. 30, Nr 4, 2041-2075 (2008), AMS, 2008
- Zhen-Qing Chen, Zhongmin Qian, Yaozhong Hu and Weian Zheng, *Stability and Approximations of Symmetric Diffusion Semigroups and Kernels*, Journal of Functional Analysis, 152, 225-280 (1998), Academic Press, 1998
- Jean-Pierre Lepeltier and Bernard Marchal, *Problème des martingales et équations différentielles stochastiques associées à un oprateur inégro-différentiel*, Annales de l' Institut Henri Poincaré (B) Vol. 12, 43-103, (1976)
- A.V. Skorokhod, *Studies in the Theory of Random Processes*, ISBN 9-780-48664240-6, Reading, Mass., Addison-Wesley Pub. Co., 1965