American Review of Mathematics and Statistics
December 2015, Vol. 3, No. 2, pp. 1-6
ISSN: 2374-2348 (Print), 2374-2356 (Online)
Copyright © The Author(s). All Rights Reserved.
Published by American Research Institute for Policy Development
DOI: 10.15640/arms.v3n2a1
URL: http://dx.doi.org/10.15640/arms.v3n2a1

Covariance Operators and the Central Limit Theory for "Loop" Markov Chains

P. Gillespie[1] & S. Molchanov[2]

Abstract

This paper discusses the class of finite of Markov Chains, representing the random transition between deterministic dynamical systems. For such chains, the covariance operator B usually has a very rich kernel. We'll give a complete analysis of B and discuss its applications to the random number generators (RNG's).

Keyword: Central Limit Theory for Markov Chains; Covariance Operator; Loop Markov Chains.

1. Introduction

The goal of this paper is to discuss the analysis of the covariance operator for the special class of Markov Chains, called *"Loop" Markov Chains*. These chains have the following description: there are several pure deterministic dynamical system and the random process jumps with some distribution from one system to another at some moment of time (which can be random or deterministic). A typical example of such a situation gives the algorithmical random number generator's (RNG's)(i.e. deterministic dynamical system with discrete time) which randomly changes states at the random moments due to seeding by ideal physical RNG's. Such *"hybrid"* RNG's have much better statistical properties than the algorithmical RNG's (see [4] and [5]).

The idea of "Loop" Markov Chain (LMC) was first proposed by Kai Lai Chung [1] in a completely different setting. We'll study the Central Limit Theorem (CLT) and the covariance operator for such chains. Our model will contain countably many loops; the case of finitely many loops is much simpler. In contrast to the case of finite Markov Chains, the D$\ddot{o}$eblin condition is not fulfilled in the case when the length of the loops are unbounded.

Description of the model: Let's consider a sequence of integers $2 \leq n_1 \leq n_2 \leq \cdots \leq n_j \leq \cdots$, where $\gcd(n_1, n_2, \cdots) = 1, n_j \to \infty \ as \ j \to \infty$. The phase space of LMC has the following structure: there is a common point O that connects all of the loops, each loop l_j consists of successive points.

$$l_j = (0, 1_j, 2_j, \cdots, n_j - 1), j = 1,2, \cdots, k, \cdots \quad (1)$$

Transition probabilities of "Loop" Markov Chains have a simple structure:

a. $p(0, 1_j) = p_j > 0, \ j = 1, 2, \cdots, k, \cdots \quad \sum_{j=1}^{\infty} p_j = 1$
b. Inside each loop, the motion is deterministic: $p(1_j, 2_j) = 1, \cdots, p(n_j - 1, 0) = 1, \quad j = 1, 2, \cdots, k$

This chain is recurrent, aperiodic, and connected (this is due to arithmetic condition on $(n_1, n_2, \cdots, n_j, \cdots)$).

The ergodicity depends on the existence of the expectation for the random variable τ_0: the first return time from $0 \to 0$. One should see that τ_0 has values $n_1, n_2, \cdots, n_j, \cdots$ with probabilities $p_1, p_2, \cdots, p_j, \cdots$, i.e. $E_0 \tau_0 = \sum_{j=1}^{\infty} p_j n_j$. We'll assume that $\sum_{j=1}^{\infty} p_j n_j = m_0 < \infty$.

[1] Fayetteville State University, NC, USA. Email: pgillespie@uncfsu.edu
[2] University of North Carolina at Charlotte, NC, USA. Email: smolchan@uncc.edu, Author was supported partially by the NSF grant DMS-0405927

Lemma 1.1: *Invariant distribution from the LMC, under the condition $\sum_i p_i n_i < \infty$, is given by*

$$\pi(0) = \frac{1}{\sum_{i=1}^{\infty} p_i n_i};\ \pi(i,j) = \frac{p_i}{\sum_{i=1}^{\infty} p_i n_i} = \frac{p_i}{m_0} \qquad (2)$$

where $i = 1, 2, \cdots, k$ *and* $j = 1, 2, \cdots, n_i - 1$. *Here* $m_0 = \sum_{i=1}^{\infty} p_i n_i = E_0 \tau_0$ *and* τ_0 *is the time of the first return from* $0 \to 0$.

Proof. Let $p_1, p_2, \cdots, p_k, \cdots$ be the probability of entering a certain loop. The $\gcd(n_1, n_2, \cdots) = 1$; therefore, the chain is aperiodic. Obviously, $\pi(i,j) = \pi(0)p_i$ and

$$\sum_{i,x} \pi_i(x) = 1 \quad \Rightarrow \quad \pi(0) + \sum_{i=1}^{\infty} \pi(0) p_i (n_i - 1) = 1$$

Finally

$$\pi(0) = \frac{1}{1 + \sum_{i=1}^{\infty} p_i (n_i - 1)} = \frac{1}{1 + \sum_{i=1}^{\infty} (p_i n_i - p_i)} = \frac{1}{1 + \sum_{i=1}^{\infty} p_i n_i - 1} = \frac{1}{\sum_{i=1}^{\infty} p_i n_i}$$

$\Longrightarrow \pi(0) = \frac{1}{\sum_{i=1}^{\infty} p_i n_i}$ and $\pi_i = \frac{p_j}{\sum_{i=1}^{\infty} p_i n_i}$ ■

For aperiodic connected Markov Chains, the ergodicity is equivalent to the positive recurrence, that is, finiteness of the expected values of the first return to any state, say 0.

2 Covariance Operators

Let's recall the Central Limit Theorem (CLT) for the Markov Chains. If $x_t,\ t = 1, 2, \cdots$ is a ergodic aperiodic Markov Chain on the countable phase space X; π is the invariant distribution of $P = [p(x,x)]$ is the transition matrix, then one can introduce the Hilbert space

$$L^2(X,\pi) = \left\{ f \epsilon X \to R^1 \vdots \sum_{y \in X} f(y)\pi(y) = 0,\ \ \sum_{y \in X} |f^2(y)| \pi(y) = \|f\|_\pi^2 < \infty \right\} \qquad (3)$$

Now let P^* be the operator, conjugated to P in $L^2(X,\pi)$. It means that $\forall \left(f_1, f_2 \in L^2(X,\pi)\right), (Pf_1 \cdot f_2)_\pi = (f_1 P^* \cdot f_2)_\pi$. It's easy to see

$$P^* = \frac{\pi(y) P(y,x)}{\pi(x)} \qquad (4)$$

where $P^*$is a stochastic matrix. Let $x_t^*,\ t = 0, 1, \cdots$ be a Markov Chain with transition probabilities P^*. The invariant measure of x_t^* is again π.

Let's formulate the Central Limit Theorem based on the theory of ergodic martingale-difference.

Theorem 2.1 *Assume that* $f \in L^2(X,\pi)$ *and the homological equation* $f = g - Pg$ *has a solution* $g \in L^2(X,\pi)$ *(formally,* $g = (I-P)^{-1} f = \sum_{k=0}^{\infty} P^k f$, *but in general this series converges* $\forall \left(f \in L^2(X,\pi)\right)$ *only under the Döeblin condition, which is not true under the condition* $n_j \to \infty$*). Then*

$$\frac{S_n}{\sqrt{n}} = \frac{\sum_{t=0}^{n-1} f(x_t)}{\sqrt{n}} \xrightarrow{law} N(0, \sigma^2) \qquad (5)$$

where

$$\sigma^2(f) = (g \cdot g)_\pi - (Pg \cdot Pg)_\pi = (Bf \cdot f)_\pi \qquad (6)$$

(limiting variance)

The covariance operator *B* (or corresponding quadratic form $(Bf \cdot f)_\pi = \sigma^2(f)$ can be presented formally as

$$\sigma^2(f) = [(g - Pg) \cdot (g + Pg)]_\pi = [f \cdot (f + 2Pf + 2P^2 f + \cdots)]_\pi = [f \cdot (f + Pf + P^* f + P^2 f + (P^*)^2 f + \cdots \pi \qquad (7)$$

$$= (f \cdot (F + F^* - I) f)_\pi \qquad (8)$$

Here $F = I + P + P^2 + \cdots$ is called the fundamental matrix of the Markov Chain $x_t,\ t = 0, 1, \cdots$. Operator F is bounded under Döeblin condition but in general it is unbounded. This is also true for the covariance operator $B = F + F^* - I$.

3 Spectral Analyses

The spectral analysis of the covariance operator B is an interesting problem with potential statistical applications. If instead of the chain $x(t)$, we can observe only the additive functional $S_n = \sum_{s=0}^{n-1} f(x_s)$ then (for fixed $L^2 -$ norm$\|f\|_\pi = 1$) the biggest amount of the information about the chain will provide the top eigenfunction Ψ_0 of B: $\lambda_0\Psi_0 = B\Psi_0$, $\lambda_0 = \max \lambda_i(B)$ (it is the leading factor in the statistical analysis). Also, the function $\Psi \in \ker B = \{\Psi: B\Psi = 0\}$ are also interesting for such functions as S_n which are bounded in probability if $n \to \infty$. The following theorem gives the complete characteristics of *ker B*.

Theorem 3.*1* *In the situation in section 2,* $\ker B = Span\,(f: f = g - Pg)$ *and* $g = P^*Pg$, *i.e.* $g \in \ker(I - P^*P)$. *Let's stress that* $I - P^*P$ *is symmetric and nonnegative on* $L^2(x, \pi)$ *due to the fact that* P^*P *is symmetric and stochastic (with top eigenvalues of 1).*

Proof. Assume that $\sigma^2(f) = 0 \Longrightarrow f = g - Pg$ and $(g \cdot g)_\pi - (Pg, Pg)_\pi = 0$, i.e. $(g, g)_\pi = (g, P^*Pg) \Rightarrow (g, (I - P^*P)g) = 0$. But $(I - P^*P)$ is symmetric and nonnegative and the spectral theorem for $(I - P^*P)$ gives immediately that $g \,\epsilon \ker(I - P^*P)$. ∎

Let's apply this general result to the LMC.

In the future, $f(x), g(x)$ will be the generic functions from $L^2(X, \pi)$, while f_0, f_{ij}, g_0, g_{ij} will be the particular function from $L^2(X, \pi)$. *F* will be the functions constant on the loops and *G* are functions linear on the loops.

Theorem 3.2 $\ker B = span\big(f_0;\ f_{ij}, i = 1, 2, \cdots, k, \cdots;\ j = 1, 2, \cdots n_i - 2\big)$. *If we have a finitely many, say k, loops of length* $n_j - 1, j = 1, 2, \cdots, k$ *then*

$$\dim \ker B = \sum_{i=1}^{k} n_i - 2k + 1\,.$$

Here

$$g_0(y) = \delta_0(y) = \begin{cases} 1 & \text{if} \quad y = 0 \\ 0 & \text{if} \quad y \neq 0 \end{cases} \qquad\qquad f_0(y) = g_0 - Pg_0 = \begin{cases} 1 & \text{if} \quad x = 0 \\ -p_i & \text{if} \quad x = (i, n_i - 1) \\ 0 & \text{if} \quad \text{otherwise} \end{cases}$$

For $1 \le j \le n_i - 2,\ \ i = 1, 2, \cdots, k$

$$g_{ij}(y) = \delta_{ij}(y), \qquad f_0(y) = g_0 - Pg_0 = \begin{cases} -1 & \text{if} \quad x = 0 \\ 1 & \text{if} \quad x = (i, j) \\ 0 & \text{if} \quad \text{otherwise} \end{cases}$$

Proof. Let's calculate the stochastic matrix $Q = P^*P$ which corresponds to the following Markov Chain $x^\pm(t)$. Assume that the initial chain $x(t)$ starts from $x \in X$. Let's make one step with transition probabilities $p^*(y, z)$ (negative direction at that time) and after one step with probabilities $p(x, y)$ (positive direction at that time). These two steps give the single transition of $x^\pm$ from *x* to *z* with probability $q(x, y) = \sum_{z \in X} p^*(x, z)p(z, y)$. Direct calculations using formula (4) gives

$$p^*\big((i, l), (i, l - 1)\big) = 1 \qquad l = 2, \cdots, n_i,\ \ i = 1, 2, \cdots, k, \cdots;$$
$$p^*\big((i, 1), 0\big) = 1 \qquad i = 1, 2, \cdots, k, \cdots;$$
$$p^*\big(0, (i, n_i - 1)\big) = p_i \qquad i = 1, 2, \cdots, k, \cdots$$

Now from the formula $q(x, y) = \sum_{z \in X} p^*(x, z)p(z, y)$ we get

$$q\big((i, l), (i, l)\big) = 1 \qquad l = 2, \cdots, n_i - 1,\ \ i = 1, 2, \cdots, k, \cdots;$$

$$q\big((i,l),(j,1)\big) = p_j \quad i,j = 1,2,\cdots,k,\cdots;$$
$$q(0,0) = 1$$

It means that for the chain $x^{\pm}(t)$ with transition matrix Q state 0; $(i,l), i = 1,2,\cdots,k,\cdots; l = 1,2,\cdots,n_i - 2$ are absorbing ones and the remaining states set $\Gamma = \{(i,n_i-1), i = 1,2,\ \cdots,k\}$ are the transient ones. ■

Remark 3.3 *We understand $span$ in theorem 3.2 as a set of the finite linear combinations of the functions f_0 and f_{ij} but not as a closed linear subspace in $L^2(X,\pi)$, except in the case when the number of loops in finite. It's easy to see if we consider function $f = \sum_{i=1}^{\infty} f_{i1}\cdot h_i$ with sufficiently fast increasing magnitudes h_i, we'll get $P-a.s.$ unbounding sums $\sum_{s=0}^{n-1} f_i h_i = S$, though formally $f \in ker\,B$.*

It is well known that all Q-harmonic functions, i.e., the solutions of the equation $g = Qg$, which can also be written as $(I-P^*P)g = 0$, are constant on the ergodic classes of the chain and $dim\{g: g = Qg\}$ is equal to the number of such classes. As a result,

$$\dim\ker(I-P^*P) = 1 + \sum_{i=1}^{k}(n_i-2) = \sum_{i=1}^{k} n_i - 2k + 1$$

For each function g_0 and $g_{(ij)}$, one can calculate the corresponding f such that $\sigma^2(f) = 0$. Namely,

$$f_0(y) = \big(g_0 - Pg_0(y)\big) = \begin{cases} 1 & \text{if} \quad y = 0 \\ -1 & \text{if} \quad y = (i,n_i-1),\ \ i = 1,2,\cdots,k \end{cases}$$
$$f_{(i,j)}(y) = \begin{cases} 1 & \text{if} \quad y = (i,j) \\ -1 & \text{if} \quad y = (i,j-1),\ \ j = 2,\cdots,n_i-1,\ \ i = i,\cdots,k \end{cases}$$

Our goal now is to give the description of $Im\,B = (\ker B)^{\perp}$ in $L^2(X,\pi)$, containing the functions $f: \sigma(f) = (Bf,f)_\pi > 0$.

Theorem 3.4 *$Im\,B \in L^2(X,\pi)$. It consist of the functions Γ, constant on the loops*

$$F(x) = \begin{cases} \varphi_i & \textit{if} \quad x = (i,l),\ \ l = 1,2,\cdots,n_i-1,\ \ i = 1,2,\cdots \\ F(0) & \textit{if} \quad x = 0 \end{cases}$$

And restricted by the linear conditions

$$F(0) = \sum_{i=1}^{\infty} p_i\,\varphi_i, \sum_{i=1}^{\infty} p_i n_i \varphi_i = 0,\ \ \sum_{i=1}^{\infty} \frac{\varphi_i^2 (n_i-1)p_i}{m(0)} + \frac{F(0)}{m(0)} < \infty.$$

In case of k loops, $dim\,Im\,B = k-1$.

Proof. Condition of orthogonality, $(F\cdot f_{(i,l)})_\pi = 0$ gives $F(i,l-1) = F(i,l);\ l = 2,\cdots,n_i-1$. The additional condition of orthogonality $(F\cdot f_0)_\pi = 0$ provides the first relation $F(0) = \sum_{i=1}^{\infty} p_i\varphi_i$. Finally, due to the definition of $L^2(X,\pi), \sum F(x)\pi(x) = 0$, i.e. $F(0) + \sum_{i=1}^{\infty}(n_i-1)p_i\varphi_i = 0$. Taking into account the previous relation, we'll get $\sum_{i=1}^{\infty} p_i\varphi_i\, n_i = 0$. ■

For each function $F \in (\ker B)^{\perp}$, we can solve in $L^2(X,\pi)$ the homological equation, $F = G - PG$ and calculate $\sigma^2(F) = (BF,F)_\pi = (G,G)_\pi - (PG,PG)_\pi$. Let

$$G_i = G(i,1)$$

then

$$G - PG = F$$

gives

$G(i,1) - G(i,2) = \varphi_i, G(i,2) - G(i,3) = \varphi_i, \cdots \Rightarrow G(i,l) = G_i - (l-1)\varphi_i,\ i = 1,2,\cdots,n_i - 1.$

Since

$$G(i,n_i-1) - G(0) = \varphi_i$$

we have

$$G(0) = G_i - (n_i-1)\varphi_i$$

i.e.

$$G_i = G(0) + (n_i - 1)\varphi_i.$$

It is easy to see the condition, $G \in L^2(X, \pi)$ is equivalent to

$$\sum_i p_i \sum_{j=1}^{n_i - 1} n_j^2 \varphi_j^2 < \infty \Rightarrow \sum_{i=1}^{\infty} p_i n_j^3 \varphi_j^2 < \infty$$

Since $G(x) \in L^2(x, \pi)$ for $G(0)$, we have the equation

$$G(0) + \sum_{i=1}^{\infty} p_i \sum_{j=1}^{n_i - 1} (G(0) + j\varphi_i) = 0,$$

i.e.

$$G(0) + \sum_{i=1}^{\infty} p_i G(0)(n_i - 1) + \sum_{i=1}^{\infty} \varphi_i p_i \frac{n_i(n_i - 1)}{2} = 0,$$

$$G(0)m(0) + \sum_{i=1}^{\infty} \varphi_i p_i \frac{n_i(n_i - 1)}{2},$$

$$G(0) = -\frac{1}{m(0)} \sum_{i=1}^{\infty} \frac{\varphi_i p_i n_i^2}{2}$$

In the last, we used the relation $\sum_{i=1}^{\infty} \varphi_i p_i n_i = 0$.

For $\sigma^2(F)$, we have

$$\sigma^2(F) = (G, G)_\pi - (PG, PG)_\pi$$

$$= \frac{G^2(0) - (G(0) - F(0))^2}{m(0)} + \frac{1}{m(0)} \sum_{i=1}^{\infty} p_i \sum_{l=1}^{n_i - 1} (G(0) + l\varphi_i)^2 - (G(0) + (l - 1)\varphi_i)^2$$

$$= \frac{2G(0)F(0) - F^2(0)}{m(0)} + \frac{1}{m(0)} \sum_{i=1}^{\infty} p_i \sum_{i=1}^{\infty} \varphi_i (2G(0) + (2l - 1)\,\varphi_i)$$

$$= \frac{2G(0)F(0) - F^2(0)}{m(0)} + \frac{1}{m(0)} \sum_{i=1}^{\infty} p_i\, \varphi_i (2G(0)(n_i - 1) + \varphi_i (n_i - 1)^2)$$

$$= \frac{2G(0)F(0) - F^2(0)}{m(0)} - \frac{2G(0)F(0)}{m(0)} + \frac{1}{m(0)} \sum_{i=1}^{\infty} p_i\, \varphi_i^2 (n_i - 1)^2$$

$$= \frac{1}{m(0)} \left[-\left(\sum_{i=1}^{\infty} \varphi_i p_i \right)^2 + \sum_{i=1}^{\infty} p_i\, \varphi_i^2 (n_i - 1)^2 \right]$$

It is better to present the final answer in the different form (using the relation $\sum_{i=1}^{\infty} p_i \varphi_i\, n_i = 0$):

$$\sigma^2(F) = \frac{1}{m(0)} \left[\sum_{i=1}^{\infty} p_i\, \varphi_i^2 (n_i - 1)^2 - \left(\sum_{i=1}^{k} \varphi_i p_i (n_i - 1) \right)^2 \right]$$

$$= \frac{1}{m(0)} \sum_{i=1}^{k} p_i [l_i(n_i - 1) - l]^2, \ \ where\ l = \sum_{i=1}^{\infty} \varphi_i p_i (n_i - 1) = -F(0)$$

One can see now that $\sigma^2(F) = 0$ iff $\varphi_i = \frac{F(0)}{n_i - 1}$, $i = 1, 2, \cdots, k, \cdots$. Vector $\overrightarrow{\varphi_0} = \left[-1, \frac{1}{n_i-1}, \cdots, \frac{1}{n_k-1}, \cdots\right]$ is an eigenvector for the covariance matrix *B* with eigenvalue $\lambda_0 = 0$. The orthogonal complement to this vector if given by

$$(F, \varphi_0)_\pi = 0 \Rightarrow -F(0) + \sum_{i=1}^{\infty} p_i \frac{\varphi_i (n_i - 1)}{n_i - 1} \Rightarrow F(0) = \sum_{i=1}^{\infty} p_i \varphi_i$$

In other terms, under conditions $F(0) = \sum_{i=1}^{\infty} p_i \varphi_i$, $\sigma^2(F) > 0$. If the number of loops is equal to $k < \infty$, this implies that $Rank\ B = k - 1$.

Reference

Kai Lai Chung (1960), *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, New York, New York.

Kemeny, J. G., Snell J. L. (1960), *Finite Markov Chains*, Van Nostrand Reinhold, Princeton, N.J.

Billingsley, Patrick (1968), *Convergence of Probability Measure*, J. Wiley, New York.

A. Figotin, A. Gordon, S. Molchanov, J. Quinn, and N. Stavrakas, *Occupancy Numbers in Testing Random Number Generators*, SIAM J. Appl. Math. 62 (2002), no. 6.

A. Gordon, S. Molchanov, and J. Quinn, *Fermi – Dirac Generators and Tests for Randomness*, CRM Proceedings Lecture Notes (American Mathematical Society), Probability and Mathematical Physics, Volume 42 (2007).