Elect. Comm. in Probab. **12** (2007), 120–133

ELECTRONIC COMMUNICATIONS in PROBABILITY

ASYMPTOTIC VARIANCE OF FUNCTIONALS OF DISCRETE-TIME MARKOV CHAINS VIA THE DRAZIN INVERSE

DAN J. SPITZNER
Department of Statistics (0439), Virginia Tech, Blacksburg, VA 24061, USA.
email: `dspitzne@vt.edu`

THOMAS R. BOUCHER
Department of Mathematics, Plymouth State University, Plymouth, NH 03264, USA.
email: `tboucher1@plymouth.edu`

Submitted August 30 2006, *accepted in final form* April 9 2007

AMS 2000 Subject classification:
Keywords: General state space Markov chains; f-regularity; Markov chain central limit theorem; Drazin inverse; fundamental matrix; asymptotic variance

Abstract

We consider a ψ-irreducible, discrete-time Markov chain on a general state space with transition kernel P. Under suitable conditions on the chain, kernels can be treated as bounded linear operators between spaces of functions or measures and the Drazin inverse of the kernel operator $I-P$ exists. The Drazin inverse provides a unifying framework for objects governing the chain. This framework is applied to derive a computational technique for the asymptotic variance in the central limit theorems of univariate and higher-order partial sums. Higher-order partial sums are treated as univariate sums on a 'sliding-window' chain. Our results are demonstrated on a simple AR(1) model and suggest a potential for computational simplification.

1 Introduction

Assume $\{\Phi\}$ is a ψ-irreducible, discrete-time Markov chain with transition kernel P evolving on a Polish state space $\mathcal{X}$. Denote by $\mathcal{B}(\mathcal{X})$ the Borel sets associated with $\mathcal{X}$. A "univariate partial sum" will refer to $S_n(g) = \sum_{i=1}^n g(\Phi_i)$ for a function $g : \mathcal{X} \to \mathbb{R}$. Define the "sliding-window chain" as $\{\tilde{\Phi}_i = (\Phi_i, \ldots, \Phi_{i+p-1})\}$ for some integer p. A "high-order partial sum" is $S_n(\tilde{g}) = \sum_{i=1}^n \tilde{g}(\Phi_i, \ldots, \Phi_{i+p-1})$ for some function $\tilde{g} : \mathcal{X}^p \to \mathbb{R}$.
We consider the role of linear operators in formulating the limiting behavior of the chain. Of special interest are the "fundamental kernel," Z, and "Drazin inverse" (or "group inverse") $Q^{\#}$, of the kernel $Q = I - P$. Our study is somewhat unconventional since it focuses on $Q^{\#}$ rather than Z, which has traditionally been used in investigations of a Markov chain's limiting behavior. The framework centering on $Q^{\#}$ is particularly tidy and convenient, and can simplify computation. As an example, we show $Q^{\#}$ allows an easy correspondence between the objects governing sliding-window chains and their analogues on the original chain.

A considerable literature is available on central limit theorems for univariate partial sums, with a variety of formulas for the asymptotic variance of $S_n(g)$. The sliding-window chain inherits the stability properties of $\{\Phi\}$ and admits a viewpoint in which $S_n(\tilde{g})$ is treated as a *univariate* partial sum, $\sum_{i=1}^n \tilde{g}(\tilde{\Phi}_i)$. Connections between the limiting properties of $S_n(\tilde{g})$ and $S_n(g)$ are made by describing the probability laws governing $\{\tilde{\Phi}\}$ in terms of those of $\{\Phi\}$.

Our investigation builds upon the central limit theorem derived in Meyn and Tweedie (1993, Ch. 17) and related ideas in Glynn and Meyn (1996). Our main reference on the Drazin inverse of a linear operator is Koliha (1996a), but Campbell and Meyer's (1979) investigations in finite-dimensional state space also provide inspiration. References for Markov-chain central limit theorems abound, including Kemeny and Snell (1960), Cogburn (1972), Nummelin (1984), Kipnis and Varadhan (1986), Tóth (1986), Geyer (1992), Chan and Geyer (1994), Tierney (1994), Robert (1995), Robert and Casella (1999, Ch. 4), Roberts and Tweedie (1996), Nummelin (2002), Hobart *et al.* (2002), Roberts and Rosenthal (2004), and Jones (2004).

The main results are in Section 2. Markov chain stability conditions and solutions to the "Poisson equation" are discussed in Section 2.1; these establish the fundamental kernel and Drazin inverse as bounded linear operators on spaces of functions and signed measures. Section 2.2 expresses representations for asymptotic variance associated with Meyn and Tweedie's (1993) central limit theorem in terms of the Drazin inverse. Extensions of the Drazin inverse methodology to the sliding window chain are in Sections 2.3 and 2.4. Demonstrations on linear autoregressive processes are in Section 3 and conclusions are made in Section 4.

2 Results

Let $\mathcal{M}$ denote the space of bounded signed measures on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. Let $\nu \in \mathcal{M}$, $g : \mathcal{X} \to \mathbb{R}$ a measurable function, A and B kernels, $x \in \mathcal{X}$, and $E \in \mathcal{B}(\mathcal{X})$. We have the following notation: $\nu A(E) := \int A(x,E)\nu(dx)$; $Ag(x) := \int g(y)A(x,dy)$; $BA(x,E) := \int A(y,E)B(x,dy)$; $\nu g := \int g(x)\nu(dx)$; $I(x,E) := 1$ if $x \in E$ and 0 otherwise; $I_g(x,E) := g(x)I(x,E)$. Define the n-step transition probability kernel P^n of $\{\Phi\}$ as the n-fold product of P with itself.

2.1 The Poisson equation and linear operator solutions

Define the kernel $R = \sum_{i=0}^{\infty} q(i)P^i$, where q is a probability measure on the nonnegative integers. A set C is *petite* if it holds that $R(x,E) \geq \nu(E)$ for every $x \in C$, $E \in \mathcal{B}(\mathcal{X})$, some q as above, and some nontrivial measure ν on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$.

A rather weak stability condition for $\{\Phi_i\}$ which yields a central limit theorem is

$$PV(x) \leq V(x) - f(x) + bI(x,C) \text{ for every } x \in \mathcal{X}, \tag{1}$$

where b is a finite constant, $f : \mathcal{X} \to [1,\infty)$, $V : \mathcal{X} \to [1,\infty)$ and $C \in \mathcal{B}(\mathcal{X})$ is a petite set. Such a chain is termed *f-regular* and has a stationary distribution π with $\pi f < \infty$ (see e.g. Glynn and Meyn, Thm. 2.2). Define the kernel $\Pi(x,E) := \pi(E)$ and note it satisfies $\Pi P = P\Pi = \Pi$.

For a functional $g : \mathcal{X} \to \mathbb{R}$ the limiting properties of $S_n(g)$ can be understood through a solution $\hat{g} : \mathcal{X} \to \mathbb{R}$ to the "Poisson equation,"

$$g - \pi g \quad = \quad \hat{g} - P\hat{g}. \tag{2}$$

When (1) holds and $\pi V < \infty$ it is known solutions to the Poisson equation exist. Moreover, these solutions may be expressed in terms of kernel operators. Glynn and Meyn (1996, Thm.

2.3) established that when (1) holds the *fundamental kernel* $Z = \{I - P + \Pi\}^{-1}$ provides solutions $\hat{g} = Zg$ to the Poisson equation (2). Meyn and Tweedie (1993, eq. 17.39) give solutions to the Poisson equation through $\hat{g} = \sum_{i \geq 0}\{P^i g - \pi g\}$. Define the kernel $Q^\# := \sum_{i \geq 0}\{P^i - \Pi\}$ and then $\hat{g} = Q^\# g$ is a solution to the Poisson equation. Thus, both Z and $Q^\#$ are kernels operating on a space of functions (defined below) to yield solutions to the Poisson equation.
In a similar spirit, we formalize a framework for operators A that send g to its solutions $\hat{g} = Ag$ of the Poisson equation (2). Following convention, for a function $h : \mathcal{X} \to [1, \infty)$, define the linear space L_h^∞ as all measurable functions $g : \mathcal{X} \to \mathbb{R}$ for which $g(x)|/h(x)$ is bounded over $x \in \mathcal{X}$. The space L_h^∞ can be equipped with the norm $\|g\|_h = \sup_x |g(x)|/h(x)$, making it a Banach space. The following is essentially Theorem 17.4.2 of Meyn and Tweedie (1993):

Proposition 1. *Suppose (1) is satisfied. If $\pi V < \infty$, then there is an $a \in \mathbb{R}$ such for each $|g| < f$ the Poisson equation has a solution $\hat{g}$ for which $|\hat{g}| \leq a(V + 1)$.*

Following Glynn and Meyn (1996, Thm. 2.3), Proposition 1 implies a linear operator which gives solutions to the Poisson equation is a bounded linear operator from L_f^∞ to L_{V+1}^∞. When the operator is a kernel, it may also be defined on linear spaces of signed measures. Define $\mathcal{M}_h$ as the space of signed measures $\nu \in \mathcal{M}$ for which $|\nu g|$ is bounded over $g \in L_h^\infty$. $\mathcal{M}_h$ is a Banach space under the norm $\|\nu\|_h = \sup_{g:|g| \leq h} |\nu g|$.

Proposition 2. *Suppose* (1) *is satisfied and $\pi V < \infty$. The kernels* Π*, $Q^\#$ and Z are bounded linear operators from $\mathcal{M}_f$ to $\mathcal{M}_{V+1}$.*

Proof. Suppose the kernel A is such that $\hat{g} = Ag$ satisfies the Poisson equation (2) for $g \in L_f$. Choose $g_1 \in L_f^\infty$, $\nu \in \mathcal{M}_{V+1}$ arbitrarily and set $g_2 = Ag_1$. Then Proposition 1 implies $g_2 \in L_{V+1}^\infty$, in turn implying $|\nu g_2| < \infty$. Since $\nu g_2 = \{\nu A\} g_1$, then $\nu A \in \mathcal{M}_f$. Thus A is a bounded linear operator from $\mathcal{M}_f$ to $\mathcal{M}_{V+1}$ and the conclusion regarding $Q^\#, Z$ follows. Since $\pi V < \infty$ then clearly $\pi \in \mathcal{M}_{V+1}$. Since (1) holds this implies $\pi f < \infty$ and we have $\nu \Pi g_1 < \infty$, implying $\nu \Pi \in \mathcal{M}_f$. Thus Π is a bounded linear operator from $\mathcal{M}_f$ to $\mathcal{M}_{V+1}$. $\square$

For a kernel A to be the *Drazin inverse* of the kernel $Q = I - P$ it must satisfy the algebraic criteria

$$QA = AQ, \quad QAQ = Q, \quad \text{and } AQA = A, \tag{3}$$

in which case it is unique in that respect. The next Proposition establishes $Q^\#$ is the Drazin inverse of Q and delineates what will be useful relationships among Z, $Q^\#$, and Π.

Proposition 3. *Suppose* (1) *holds and $\pi V < \infty$. Consider the kernels* Π*, $Q^\#$ and Z defined as operators either from $\mathcal{M}_{V+1}$ to $\mathcal{M}_f$ or L_f^∞ to L_{V+1}^∞.*

(i) $Q^\#$ *satisfies the criteria (3) and is therefore the unique Drazin inverse of Q,*

(ii) $Q^\# = Z\{I - \Pi\} = \{I - \Pi\}Z$,

(iii) $\Pi = I - Q^\# Q = I - QQ^\#$,

(iv) $Z = I + PQ^\#$, *and*

(v) $Q^\# = P^j Q^\# + \sum_{i=0}^{j-1}\{P^i - \Pi\}$.

Proof. Note when (1) is true and $\pi V < \infty$, the kernels Π, $Q^{\#}$ and Z exist and are well-defined due to Propositions 1 and 2.
To prove statement (*i*), it is straightforward to use the definitions of Q and $Q^{\#}$ to show (3). The uniqueness of $Q^{\#}$ with respect to the criteria (3) follows from Koliha (1996, lem. 2.4).
Statement (*ii*) follows from $\{I-\Pi\}$ commuting with Z^{-1}: $\{I-\Pi\}Z = ZZ^{-1}\{I-\Pi\}Z = Z\{I-\Pi\}Z^{-1}Z = Z\{I-\Pi\}$.
Statement (*iii*) now follows from (*ii*), recalling $Q = I - P$.
Statement (*iv*) also follows immediately from (*ii*) since $Z = \{I - P + \Pi\}^{-1}$ and

$$\{I + PQ^{\#}\}Z^{-1} \quad = \quad Z^{-1} + P\{I-\Pi\}ZZ^{-1} = I.$$

Statement (*v*) for the case $j = 1$ is immediate from statement (*iii*), since $I - \Pi = QQ^{\#}$ is equivalent to $Q^{\#} = PQ^{\#} + \{I-\Pi\}$. To prove the general case, apply this in the induction step

$$Q^{\#} = P^{j-1}Q^{\#} + \sum_{i=0}^{j-2}\{P^i - \Pi\} = P^{j-1}[PQ^{\#} + \{I-\Pi\}] + \sum_{i=0}^{j-2}\{P^i - \Pi\} = P^jQ^{\#} + \sum_{i=0}^{j-1}\{P^i - \Pi\}.$$

□

The results to this point, and those to follow, require the drift condition (1) as a sufficient condition. We pause for a remark on the necessity of (1). The existence of the operator $Q^{\#}$ follows from f-regularity with $\pi(V) < \infty$, since for a ψ-irreducible chain these imply $\sum_{i\geq 0}|P^i g - \pi g| < \infty$ for any initial x and any function g with $|g| \leq f$ (Meyn and Tweedie, 1993, Thm. 14.0.1). This is not equivalent to f-ergodicity, which merely has $\| P^n - \pi \|_f \rightarrow 0$. Of course, if a chain is not f-ergodic there is no hope of $Q^{\#}$ existing, so that an investigation into the necessity of (1) must involve f-ergodic chains. A ψ-irreducible chain is f-regular if and only if it satisfies (1) (Meyn and Tweedie, 1993, Thm. 14.2.6), and, while an aperiodic f-regular chain is f-ergodic, aperiodic f-ergodic chains need not be f-regular (Meyn and Tweedie, 1993, Thm. 14.3.3). However, the f-ergodic chain restricted to a full (a set A with $\psi(A) = 1$) absorbing set is f-regular. Thus, when the maximal irreducibility measure ψ is Lebesgue measure, an aperiodic f-ergodic chain is f-regular almost everywhere. In this case, which we suspect will contain most real applications, the drift condition (1) is 'almost' necessary as well as being sufficient for the existence of $Q^{\#}$ when $\pi(V) < \infty$.
The question of the necessity of the drift condition (1) in establishing the existence of the operator Z is resolved similarly, with the difference being Glynn and Meyn (1993, Thm. 2.3) require for any $x \in \mathcal{X}$ that $E_x[\sum_{k=0}^{\tau_B - 1} f(\Phi_k)] < V(x) + c(B)$, with V as above, τ_B being the hitting time for a set $B \in \mathcal{B}(\mathcal{X})$, and $c(\cdot)$ a set function with $c(\mathcal{B}) < \infty$. This is slightly stronger than the characterization, equivalent to (1), of a ψ-irreducible chain being f-regular when it has a countable covering of the state space $\mathcal{X}$ by *regular* sets C_i with $\sup_{x\in C_i} E_x[\sum_{k=0}^{\tau_B - 1} f(\Phi_k)] < +\infty$.
The special case of *V-uniform ergodicity* occurs when the bounding function V in (1) is a multiple of f. The operator norm $\|A\|_V = \sup_{x\in\mathcal{X}} \sup_{g:|g|\leq V} |Ag(x)|/V(x)$ makes the relevant kernels bounded linear operators from either L_V^∞ to itself or $\mathcal{M}_V$ to itself. It also allows an equivalent characterization of V-uniform ergodicity: $\|P^i - \Pi\|_V \leq c\rho^i$, $c \in (0,\infty)$, $0 < \rho < 1$, providing an immediate proof of the existence of $Q^{\#}$ since $\sum_{i\geq 0}\{P^i - \Pi\}$ converges in the norm $\|\cdot\|_V$, considered over either spaces of functions or measures.
When $Q^{\#}$ is a bounded linear operator on a Banach space, operator theory results such as those of Koliha (1996b) describe its spectral properties. One particular implication is that

Q must have an isolated eigenvalue at zero, which admits the study of "second eigenvalue" properties of Markov chains. Such ideas have been used to investigate a Markov chain's rate of convergence to its invariant distribution. See, *e.g.*, Lawler and Sokal (1988) and more recent work by François (1999) and León (2001).

2.2 Representations of asymptotic variance

Considering the kernels as operators between spaces of functions makes sense when concerned with calculating asymptotic quantities of partial sums. Considering the kernels as operators between spaces of measures is logical when concerned with convergence of measures, such as convergence of the transition probabilities to the stationary. A combination of the two is applicable to studying convergence of normalized partial sums to a normal distribution in the Central Limit theorem. We recount some important asymptotic properties of the partial sum $S_n(g)$. Statements (*i*) and (*ii*) are Theorems 14.0.1 and 17.4.4 of Meyn and Tweedie (1993), respectively.

Proposition 4. *Suppose (1) is satisfied and $\pi V < \infty$.*

(i) The asymptotic mean of $n^{-1}S_n(g)$ is

$$\mu_g \quad = \quad \pi g = \int g(x)\pi(dx). \tag{4}$$

(ii) Suppose $\hat{g}$ solves (2) and $\pi(\hat{g}^2) < \infty$. If $\sigma_g^2 = \pi(\hat{g}^2 - \{P\hat{g}\}^2) > 0$ then

$$n^{-1/2}\{S_n(g) - n\mu_g\} \to N(0, \sigma_g^2). \tag{5}$$

The quantity σ_g^2 is the asymptotic variance of the partial sum $S_n(g)$. It is a quantity of interest in Markov chain theory; for example in the CLT as in Proposition 4.*ii*, also in MCMC methodology. We express the asymptotic variance in terms of the Drazin inverse $Q^{\#}$.

Proposition 5. *When Proposition 4.*ii *holds the asymptotic variance may be written*

$$\sigma_g^2 \quad = \quad 2\pi I_g Q^{\#} g - \pi I_g\{I - \Pi\}g, \tag{6}$$

$$\sigma_g^2 \quad = \quad 2\pi I_g Z\{I - \Pi\}g - \pi I_g\{I - \Pi\}g, \tag{7}$$

$$\sigma_g^2 \quad = \quad \pi I_g\{I + P\}Q^{\#} g, \tag{8}$$

$$\sigma_g^2 \quad = \quad 2\pi I_g P^j Q^{\#} g + 2\sum_{i=0}^{j-1} \pi I_g\{P^i - \Pi\}g - \pi I_g\{I - \Pi\}g. \tag{9}$$

Proof. Write $\sigma_g^2 = \pi(\hat{g}^2 - \{P\hat{g}\}^2) = \pi(\hat{g}^2 - \{\hat{g} - \bar{g}\}^2)$. Substituting $\bar{g} = \{I - \Pi\}g$, $\hat{g} = Q^{\#}g$ and simplifying yields $\sigma_g^2 = 2\pi(\bar{g}Q^{\#}g) - \pi(\bar{g}^2) = 2\pi I_{\bar{g}}Q^{\#}g - \pi I_{\bar{g}}\{I - \Pi\}g$. Now use the identities $\pi I_{\bar{g}} = \pi I_g\{I - \Pi\}$ and $\Pi Q^{\#} = 0$ (by Proposition 3.*iii*) to establish (6). Substituting $Q^{\#} = Z\{I - \Pi\}$, as in Proposition 3.*iv*, gives an expression (7) in terms of the fundamental matrix. Using $I - \Pi = QQ^{\#}$, by Proposition 3.*ii*, one is led to the compact formula (8). Alternatively, Proposition 3.*v* leads to the expression (9) for any $j \geq 1$. □

Note (9) may be interpreted in terms of the autocovariances of $\{\Phi_i\}$. Assuming Φ_0 is initially distributed according to π, it is straightforward to show $\pi I_g\{I - \Pi\}g = Var[g(\Phi_0)]$ and

$\pi I_g\{P^i - \Pi\}g = Cov[g(\Phi_0), g(\Phi_i)]$. Under (1), $P^j \to \Pi$, and since $\Pi Q^\# = 0$ it follows the first term in (9) disappears. The resulting formula is then the familiar (Meyn and Tweedie, 1993, Thm. 17.5.3)

$$\sigma_g^2 \quad = \quad 2\sum_{i=0}^{\infty} Cov[g(\Phi_0), g(\Phi_i)] - Var[g(\Phi_0)]. \tag{10}$$

2.3 Representations for a sliding window chain in terms of the original chain

Recall $\{\tilde{\Phi}_i = (\Phi_i, \ldots, \Phi_{i+p-1})\}$ and $S_n(\tilde{g}) = \sum_{i=1}^n \tilde{g}(\tilde{\Phi}_i)$. Under suitable regularity conditions on $\{\tilde{\Phi}\}$, the asymptotic properties of $S_n(\tilde{g})$ are detailed in formulas (4) - (6) with π, P, and $Q^\#$ replaced by $\tilde{\pi}$, $\tilde{P}$, and $\tilde{Q}^\#$, analogous objects governing $\{\tilde{\Phi}_i\}$. In this section, the connection of $\tilde{\pi}$, $\tilde{P}$, and $\tilde{Q}^\#$ to P, π, and $Q^\#$ is made explicit.
The transition probability kernel $\tilde{P}$ of the sliding window chain is defined on the product space $(\mathcal{X}^p, \mathcal{B}(\mathcal{X})^p)$. We extend measures from $\{\Phi\}$ to $\{\tilde{\Phi}\}$ as follows. Define $\tilde{P}^*(x_1, E) := P(\Phi_2 \in E_2, \ldots, \Phi_p \in E_p | \Phi_1 = x_1)$. Suppose $\nu \in \mathcal{M}$. For $\tilde{E} \in \mathcal{B}(\mathcal{X})^p$, define $\tilde{\nu}(\tilde{E}) := \int_{E_1} \tilde{P}^*(x_1, E)\nu(dx_1)$; for a function $\tilde{g} : \mathcal{X}^p \to \mathbb{R}$ define $\tilde{\nu}\tilde{g}$ appropriately. Define from $\tilde{g}$ the function $\tilde{g}^* : \mathcal{X} \to \mathbb{R}$ by $\tilde{g}^* := \int \tilde{g} d\tilde{P}^*$ and note we can write $\nu\tilde{g}^* = \tilde{\nu}\tilde{g}$.
An analogue to the drift condition (1) holds for the sliding window chain. From f and V in (1) define $\tilde{f}(\tilde{x}) = f(x_p)$ and $\tilde{V}(\tilde{x}) = V(x_p)$; from the petite set C define $\tilde{C} = \{\tilde{x} \in \mathcal{X}^p : x_p \in C\}$. The stability condition (1) may now be rewritten

$$\tilde{P}\tilde{V}(\tilde{x}) \leq \tilde{V}(\tilde{x}) - \tilde{f}(\tilde{x}) + bI(\tilde{x}, \tilde{C}), \qquad \text{for every } \tilde{x} \in \mathcal{X}^p. \tag{11}$$

It remains to show that $\tilde{C}$ is petite.

Proposition 6. *If C is petite for $\{\Phi\}$, then $\tilde{C} = \{(x_1, \ldots, x_p) \in \mathcal{X}^p : x_p \in C\}$ is petite for $\{\tilde{\Phi}\}$.*

Proof. Pick $\tilde{x} \in \mathcal{X}^p$ and $\tilde{E} \in \mathcal{B}(\mathcal{X})^p$. Note $\tilde{P}^{i+p-1}(\tilde{x}, \tilde{E}) = \int_{E_1} \tilde{P}^*(x_{p+1}, E)P^i(x_p, dx_{p+1})$. Define $\tilde{R} = \sum_{i=0}^\infty \tilde{q}(i)\tilde{P}^i$, where $\tilde{q}(i) = 0$ if $i < p-1$ and $\tilde{q}(i) = q(i-p+1)$ else, so that $\tilde{R}(\tilde{x}, \tilde{E}) = \int_{E_1} \tilde{P}^*(x_{p+1}, E)R(x_p, dx_{p+1})$. Since C is petite in $\mathcal{X}$ this leads to $\tilde{R}(\tilde{x}, \tilde{E}) \geq \tilde{\nu}(\tilde{E})$ for every $\tilde{x} \in \tilde{C}$ and $\tilde{E} \in \mathcal{B}(\mathcal{X})^p$, implying $\tilde{C}$ is petite. □

We now define a number of relevant linear spaces. Let $\mathcal{M}^p$ be the space of bounded signed measures on the product space $(\mathcal{X}^p, \mathcal{B}(\mathcal{X})^p)$. Recall the spaces L_h^∞ and $\mathcal{M}_h$ defined in Section 2.1. Let $\tilde{L}_h^\infty$ denote the space of measurable functions $\tilde{g} : \mathcal{X}^p \to \mathbb{R}$ such that $\tilde{g}^* \in L_h^\infty$. Let $\tilde{\mathcal{M}}_h$ denote the space of bounded signed measures $\tilde{\nu} \in \mathcal{M}^p$ for which there is a signed measure $\nu \in \mathcal{M}_h$ such that $\tilde{\nu}\tilde{g} = \nu\tilde{g}^*$ for any $\tilde{g} \in \tilde{L}_h^\infty$.
Based on Proposition (6) and the preceding discussion we expect if the original chain is stable so is the sliding-window chain. Then by the results in Section 2.2 an invariant measure must exist, a Drazin inverse must exist and so also must solutions to the sliding-window version of the Poisson equation. The following shows this is the case.

Proposition 7. *Let $\tilde{g} : \mathcal{X}^p \to \mathbb{R}$ be a measurable function. Suppose (1) holds and $\pi V < \infty$.*

(i) The measure $\tilde{\pi} \in \tilde{\mathcal{M}}_f$ is the invariant measure of the sliding window chain.

(ii) The Drazin inverse of $\tilde{Q}$, denoted $\tilde{Q}^{\#}$, exists as a bounded linear operator from $\tilde{L}_f^\infty$ to $\tilde{L}_{V+1}^\infty$ and from $\tilde{\mathcal{M}}_f$ to $\tilde{\mathcal{M}}_{V+1}$, is unique and provides solutions $\hat{\tilde{g}} = \tilde{Q}^{\#}\tilde{g}$ to the Poisson equation.

Proof. Statement (i): If $\tilde{\nu} \in \tilde{\mathcal{M}}_f$ is such that $\tilde{\nu}\tilde{g} = \nu\tilde{g}^*$ for $\nu \in \mathcal{M}_f$, any $\tilde{g} \in \tilde{L}_f$, $\nu \in \mathcal{M}_f$, then $\tilde{\nu}\tilde{P}^i\tilde{g} = \nu P^i \tilde{g}^*$ for any $i \geq 0$. Since (11) holds an invariant measure exists and is unique. The measure $\tilde{\pi} \in \tilde{\mathcal{M}}_f$ such that $\tilde{\pi}\tilde{g} = \pi\tilde{g}^*$ with $\pi \in \mathcal{M}_f$ is therefore the invariant measure of the sliding window chain.
Statement *(ii)*: Since $\pi V < \infty$ then $\tilde{\pi}\tilde{V} < \infty$. Define $\tilde{Q} = I - \tilde{P}$. As implied by Proposition 3.i, $\tilde{Q}^{\#}$ exists, is unique, and may be written $\tilde{Q}^{\#} = \sum_{i=0}^{\infty}\{\tilde{P}^i - \tilde{\Pi}\}$. It is easily verified this provides solutions to the Poisson equation. □

Since $\tilde{\pi}\tilde{V} < \infty$, the central limit theorem and asymptotic variance expressions in Propositions 4 and 5 apply to $\{\tilde{\Phi}\}$ with $\tilde{P}, \tilde{\Pi}, \tilde{Q}^{\#}$ replacing $P, \Pi, Q^{\#}$. We have already related $\tilde{P}$ and $\tilde{\Pi}$ to P and Π. We next connect $\tilde{Q}^{\#}$ to its analogous object $Q^{\#}$ on the original chain.

Proposition 8. *For a measure $\tilde{\nu} \in \tilde{\mathcal{M}}_{V+1}$, the measure $\tilde{\nu}\tilde{Q}^{\#}$ is in $\tilde{\mathcal{M}}_f$ and acts on functions $\tilde{g} \in \tilde{L}_f^\infty$ by $\tilde{\nu}\tilde{Q}^{\#}\tilde{g} = \nu Q^{\#}\tilde{g}^*$.*

Proof. Recall by definition we extended measures $\nu \in \mathcal{M}_{V+1}$ by $\tilde{\nu}\tilde{g} = \nu\tilde{g}^*$ for any $\tilde{g} \in \mathcal{L}_{V+1}^\infty$. By Proposition 7, $\tilde{Q}^{\#}$ takes measures $\tilde{\nu} \in \tilde{\mathcal{M}}_{V+1}$ to $\tilde{\mathcal{M}}_f$. Since $\tilde{\nu}\tilde{Q}^{\#} = \widetilde{\nu Q^{\#}}$ this implies $\tilde{\nu}\tilde{Q}^{\#}\tilde{g} = \widetilde{\nu Q^{\#}}\tilde{g} = \nu Q^{\#}\tilde{g}^*$. □

2.4 Asymptotic variance of high-order partial sums

We apply these results to show how the asymptotic variance of $S_n(\tilde{g})$ can be expressed in terms of $Q^{\#}$.

Proposition 9. *When Proposition 4.*ii *holds for analogous objects of the sliding window chain, the asymptotic variance of $S_n(\tilde{g})$ is*

$$\sigma_{\tilde{g}}^2 = 2\pi P_{\tilde{g}} Q^{\#}\tilde{g}^* + 2\left\{\left(\sum_{i=0}^{p-2}\tilde{\pi} I_{\tilde{g}}\tilde{P}^i\tilde{g}\right) - (p-1)(\tilde{\pi}\tilde{g})^2\right\} - \tilde{\pi} I_{\tilde{g}}\{I - \tilde{\Pi}\}\tilde{g}, \tag{12}$$

where

$$P_{\tilde{g}}(x_1, E) = \int I(x_p, E)\tilde{g}(x_1, \ldots, x_p)P(x_{p-1}, dx_p)P(x_{p-2}, dx_{p-1})\cdots P(x_1, dx_2).$$

Proof. It follows almost directly from Proposition 8 that $\tilde{\pi} I_{\tilde{g}}\tilde{P}^{p-1}\tilde{Q}^{\#}\tilde{g} = \pi P_{\tilde{g}} Q^{\#}\tilde{g}^*$. Combine this with Proposition 3.v with $j = p-1$ and (6) in which the term $\tilde{\pi} I_{\tilde{g}}\{\tilde{P}^i - \tilde{\Pi}\}\tilde{g}$ has been expanded to $\tilde{\pi} I_{\tilde{g}}\tilde{P}^i\tilde{g} - (\tilde{\pi}\tilde{g})^2$.

□

3 Applications to autoregressive processes

At this point, we have established the Drazin inverse methodology for $\{\Phi\}$, extended it to $\{\tilde{\Phi}\}$ and expressed this extension in terms of the kernels related to $\{\Phi\}$. We will now illustrate all of this with some applications.

We consider a simple time series model, the linear autoregressive process of order one, AR(1). The AR(1) model is $\Phi_{i+1} = \phi\Phi_i + \epsilon_{i+1}$, where the $\{\epsilon_i\}$ are mean-zero Gaussian random variables with constant variance σ^2. We suppose $|\phi| < 1$ so that $\{\Phi_i\}$ is a mean zero stationary process. The transition probability kernel identifies $P(x, \cdot)$ with a Gaussian distribution having mean ϕx and variance σ^2. A straightforward iterative argument will then show that $P^i(x, \cdot)$ is a Gaussian probability measure with mean $\phi^i x$ and variance $\sigma_i^2 = \sigma^2(1-\phi^{2i})/(1-\phi^2)$. Known results (see, *e.g.,* Brockwell and Davis, 1991, Ch. 3) establish the invariant distribution π is mean-zero Gaussian with variance $\sigma_\pi^2 = \sigma^2 \sum_{i=0}^{\infty} \phi^{2i} = \sigma^2/(1-\phi^2)$.

We verify the drift condition (1) for V-uniform ergodicity (and thus f-regularity, since f is a multiple of V) using the bounding function $V(x) = e^{c|x|}$, $c > 0$, which satisfies $V \geq 1$. Note $PV(x) < \infty$ and $\{\Phi_t\}$ is ψ-irreducible with the irreducibility measure ψ being Lebesgue measure. Since $|\phi| < 1$ and $E[e^{c|\epsilon_i|}] < \infty$, for $|x|$ large enough it holds that $PV(x)/V(x) < 1$, implying there exists $\beta > 0$ so that $PV(x) + \beta V(x) \leq V(x)$. Let $C = \{x : PV(x) \geq V(x)\}$ and let $b = \max_{x \in C} V(x)$; then C is petite and the drift condition for V-uniform ergodicity is satisfied: $PV(x) - V(x) \leq -\beta V(x) + bI(x, C)$.

First, in Example 1, to demonstrate the basic techniques described in Section 2.2, we treat a problem involving a univariate partial sum only. Next, in Example 2, to demonstrate the techniques derived in Section 2.4, we treat a problem involving a high-order partial sum. A third example is also given, which combines the formulas derived in Example 2 to rederive the asymptotic variance of the least squares estimator for the autoregressive parameter.

Although the examples deduce fairly broad formulas, they have been laid out in such a way as to sketch how our high-order asymptotic formulas might be used in a general methodology: the first problem provides an understanding of the action of $Q^\#$ as applied to a univariate partial sum, which is then applied in the second, more complicated problem involving higher-order partial sums; finally, the third problem pulls the results together in a useful context. The amount of computational effort required in each example is roughly the same as that which would be needed in taking a more direct approach, such as applying (10), *e.g.*, which does not involve $Q^\#$. Nevertheless, we anticipate advantages in future applications to general autoregressive time series in which $Q^\#$ is known with uncertainty and at considerable computational cost. By following a sequence of steps similar to these examples, one avoids calculation of the Drazin inverse of the sliding window chain, thereby heading off additional uncertainty and computational costs that would arise in working with this more complicated chain directly. Further discussion of potential applications is given in Section 4.

Example 1. We calculate the asymptotic variance (6) for $g(x) = x^d$ for some integer d. By rescaling $P^i(x, \cdot)$ with respect to the distribution of a standard Gaussian random variable $Z \sim G(0, 1)$, the binomial theorem provides $(\sigma_i Z + \phi^i x)^d = \sum_{j'=0}^{d} \binom{d}{j'} \{\sigma_i Z\}^{j'} \{\phi^i x\}^{d-j'}$. This leads to the formula

$$P^i g(x) \quad = \quad \sum_{j=0}^{\lfloor d/2 \rfloor} \binom{d}{2j} \sigma^{2j} \left(\frac{1-\phi^{2i}}{1-\phi^2}\right)^j \phi^{i(d-2j)} x^{d-2j} E[Z^{2j}], \tag{13}$$

in which indexing uses $j' = 2j$ only, since the odd moments of Z are zero. Here, $\lfloor d/2 \rfloor$ represents the greatest integer not exceeding $d/2$. The binomial theorem may be used again to expand $(1-\phi^{2i})^j = \sum_{k=0}^{j} \binom{j}{k} (-1)^k \phi^{2ik}$. Recalling $Q^\# g(x) = \sum_{i=0}^{\infty} \{P^i g(x) - \pi g\}$ and the identity $\sum_{i=0}^{\infty} \phi^{i(d-2j+2k)} = (1-\phi^{d-2j+2k})^{-1}$ leads to

$$Q^{\#}g(x) = \sum_{j=0}^{\lfloor d/2 \rfloor} \sum_{k \in \mathcal{K}_j} \sigma_\pi^{2j} (-1)^k \binom{d}{2j} \binom{j}{k} \left(\frac{1}{1-\phi^{d-2j+2k}} \right) x^{d-2j} E[Z^{2j}], \tag{14}$$

where $\mathcal{K}_j = 0, \ldots, j$ if $j < d/2$ and $\mathcal{K}_j = 1, \ldots, j$ if $j = d/2$. The special indexing arises since $\pi g = \sigma_\pi^d E[Z^d]$ is identical to the expanded summand in (13) indexed by $j = d/2$ and $k = 0$. The first term in (6) involves $\pi I_g Q^{\#} g$. To evaluate this using (14), write $h(x) = x^{d-2j}$, note $\pi I_g h = \sigma_\pi^{2(d-j)} E[Z^{2(d-j)}]$ and distribute.

It then remains to calculate $\pi I_g \{I - \Pi\} g$, the second term in (6). This is easily done using

$$\pi I_g \{I - \Pi\} g \quad = \quad \pi(g^2) - (\pi g)^2 = \sigma_\pi^{2d} \{E[Z^{2d}] - E[Z^d]^2\} = \sigma_\pi^{2d} Var[Z^d].$$

Thus, the asymptotic variance (6) for $g(x) = x^d$ is

$$\sigma_g^2 = 2\sigma_\pi^{2d} \sum_{j=0}^{\lfloor d/2 \rfloor} \sum_{k \in \mathcal{K}_j} (-1)^k \binom{d}{2j} \binom{j}{k} \left(\frac{1}{1-\phi^{d-2j+2k}} \right) E[Z^{2(d-j)}] E[Z^{2j}] - \sigma_\pi^{2d} Var[Z^d].$$

Expressions for the example values $d = 1, 2, 3$ are

$$\sigma_g^2 \quad = \quad \begin{cases} \left(\dfrac{\sigma^2}{1-\phi^2} \right) \left(\dfrac{1+\phi}{1-\phi} \right) & d = 1 \\[2ex] 2 \left(\dfrac{\sigma^2}{1-\phi^2} \right)^2 \left(\dfrac{1+\phi^2}{1-\phi^2} \right) & d = 2 \\[2ex] \left(\dfrac{\sigma^2}{1-\phi^2} \right)^3 \left[30 \left(\dfrac{1}{1-\phi^3} \right) + 18 \left\{ \left(\dfrac{1}{1-\phi} \right) - \left(\dfrac{1}{1-\phi^3} \right) \right\} - 15 \right] & d = 3. \end{cases}$$

In general settings, the even moments of Z may be calculated through the well-known formula $E[Z^{2j}] = 2^j \pi^{-1/2} \Gamma(j + 1/2)$, where Γ is the gamma function.

Example 2. In this example, the sliding-window chain formulas derived in Section 2.4 are used to evaluate the asymptotic behavior of the partial sums $S_1 = \sum_{i=1}^n \Phi_i \Phi_{i+r-1}$ and $S_2 = \sum_{i=1}^n \Phi_i \Phi_{i+p-1}$, with r and p being fixed, positive integers. The specific objective is to evaluate $\mu_i \approx n^{-1} E[S_i]$, $\sigma_i^2 \approx n^{-1} Var[S_i]$, for $i = 1, 2$, and $\sigma_{12} \approx n^{-1} Cov[S_1, S_2]$ under the simple AR(1) model. The parameter r shall be restricted to $1 \leq r < (p+1)/2$, which identifies the simpler cases in calculating the second term in (12).

Set $\tilde{g}(\tilde{x}) = a_1 x_1 x_r + a_2 x_1 x_p$, for arbitrary constants a_1 and a_2. Expressions for the asymptotic mean and variance of $S_n(\tilde{g})$ will have the form $\mu_{\tilde{g}} = a_1 \mu_1 + a_2 \mu_2$ and $\sigma_{\tilde{g}}^2 = a_1^2 \sigma_1^2 + 2a_1 a_2 \sigma_{12} + a_2^2 \sigma_2^2$. Thus, once $\mu_{\tilde{g}}$ and $\sigma_{\tilde{g}}$ are derived, the desired quantities may simply be read off.

We begin the derivation of $\sigma_{\tilde{g}}^2$ by calculating the first term in (12) from Proposition 9. We need the following lemma.

Lemma 1. *Define* $\tilde{e}_i(\tilde{x}) = x_i$, $\tilde{e}_{ij}(\tilde{x}) = x_i x_j$, $\tilde{e}_{ijk}(\tilde{x}) = x_i x_j x_k$, *where* $1 < i \leq j \leq k \leq p$. *Then using the* $*$ *notation defined in Section 2.3,*

$$\begin{aligned} \tilde{e}_{i,j,k}^*(x_1) &= x_1^3 \phi^{i+j+k-3} + 3x_1 \phi^{j+k-i-1} \sigma_{i-1}^2 + x_1 \phi^{i+k-j-1} \sigma_{j-i}^2 \\ \tilde{e}_{i,j}^*(x_1) &= x_1^2 \phi^{i+j-2} + \phi^{j-i} \sigma_{i-1}^2 \\ \tilde{e}_i^*(x_1) &= x_1 \phi^{i-1}. \end{aligned}$$

Proof. The expressions follow directly from formula (13). It is helpful to note that whenever $g(\tilde{x}) = g(x_i, x_j, x_k)$ (*i.e.*, g is a function of x_i, x_j, and x_k only) then

$$\tilde{g}^*(x_1) \quad = \quad \int \tilde{g}(x_i, x_j, x_k) P^{k-j}(x_j, dx_k) P^{j-i}(x_i, dx_j) P^{i-1}(x_1, dx_i).$$

□

Now observe

$$\tilde{g}^*(x_1) \quad = \quad a_1 x_1 \tilde{e}_r^*(x_1) + a_2 x_1 \tilde{e}_p^*(x_1) = (a_1 \phi^{r-1} + a_2 \phi^{p-1}) x_1^2, \tag{15}$$

to which one may immediately apply (14) with $d = 2$ in order to evaluate

$$Q^{\#} \tilde{g}^*(x_1) \quad = \quad (a_1 \phi^{r-1} + a_2 \phi^{p-1}) \left(\frac{x_1^2 - \sigma_\pi^2}{1 - \phi^2} \right).$$

Thus,

$$\begin{aligned} P_{\tilde{g}} Q^{\#} \tilde{g}^*(x_1) &= \left(\frac{a_1 \phi^{r-1} + a_2 \phi^{p-1}}{1 - \phi^2} \right) \int \{x_p^2 - \sigma_\pi^2\} \{a_1 x_1 x_r + a_2 x_1 x_p\} P^{p-1}(x_1, dx_p) \\ &= \left(\frac{a_1 \phi^{r-1} + a_2 \phi^{p-1}}{1 - \phi^2} \right) \{a_1 x_1 \tilde{e}_{r,p,p}^*(x_1) + a_2 x_1 \tilde{e}_{p,p,p}^*(x_1) - \sigma_\pi^2 \tilde{g}^*(x_1)\}; \end{aligned}$$

hence, with (13) and some straightforward manipulation, the required formula is

$$\pi P_{\tilde{g}} Q^{\#} \tilde{g}^* = 2 a_1^2 \left(\frac{\phi^{2(p-1)}}{1 - \phi^2} \right) \sigma_\pi^4 + 2 a_1 a_2 \phi^{p-r} \left(\frac{\phi^{2(r-1)} + \phi^{2(p-1)}}{1 - \phi^2} \right) \sigma_\pi^4 + 2 a_2^2 \left(\frac{\phi^{2(p-1)}}{1 - \phi^2} \right) \sigma_\pi^4. \tag{16}$$

To calculate the second term in (12), observe that (15) immediately provides

$$\pi \tilde{g}^* \quad = \quad (a_1 \phi^{r-1} + a_2 \phi^{p-1}) \sigma_\pi^2, \tag{17}$$

so the main task is to calculate the sum $\sum_{i=0}^{p-2} \tilde{\pi} I_{\tilde{g}} \tilde{P}^i \tilde{g}$. To this end, note

$$\tilde{P}^i \tilde{g}(\tilde{x}) \quad = \quad \begin{cases} a_1 x_{i+1} x_{i+r} + a_2 \phi^i x_{i+1} x_p & \text{for } 0 \leq i < p - r \\ \\ (a_1 \phi^{i-(p-r)} + a_2 \phi^i) x_{i+1} x_p & \text{for } p - r \leq i < p - 1, \end{cases}$$

and using $\tilde{h}_i(\tilde{x}) = \{\tilde{g}(\tilde{x})\}\{\tilde{P}^i \tilde{g}(\tilde{x})\}$,

$$\tilde{h}_i(\tilde{x}) = \begin{cases} a_1^2 x_1 x_{i+1} x_r x_{i+r} + a_1 a_2 x_1 x_{i+1} x_{i+r} x_p & \\ +a_1 a_2 \phi^i x_1 x_{i+1} x_r x_p + a_2^2 \phi^i x_1 x_{i+1} x_p^2 & \text{for } 0 \leq i < p - r \\ \\ a_1^2 \phi^{i-(p-r)} x_1 x_{i+1} x_r x_p + a_1 a_2 \phi^{i-(p-r)} x_1 x_{i+1} x_p^2 & \\ +a_1 a_2 \phi^i x_1 x_{i+1} x_r x_p + a_2^2 \phi^i x_1 x_{i+1} x_p^2 & \text{for } p - r \leq i < p - 1. \end{cases}$$

The restriction $r < (p+1)/2$ delineates three cases within which Lemma 1 and (13) may be used to calculate $\pi \tilde{h}_i^*$:

Case 1, $0 \leq i < r - 1 < p - r$,

$$\begin{aligned} \pi \tilde{h}_i^* &= a_1^2 (\phi^{2i} + 2\phi^{2(r-1)}) \sigma_\pi^4 + a_1 a_2 \phi^{p-r} (1 + \phi^{2i} + 4\phi^{2(r-1)}) \sigma_\pi^4 \\ &\quad + a_2^2 (\phi^{2i} + 2\phi^{2(p-1)}) \sigma_\pi^4; \end{aligned} \tag{18}$$

Case 2, $r-1 \le i < p-r$,

$$\begin{aligned}\pi\tilde{h}_i^* &= a_1^2(2\phi^{2i}+\phi^{2(r-1)})\sigma_\pi^4 + a_1a_2\phi^{p-r}(1+2\phi^{2i}+3\phi^{2(r-1)})\sigma_\pi^4 \\ &\quad + a_2^2(\phi^{2i}+2\phi^{2(p-1)})\sigma_\pi^4;\end{aligned}$$

and Case 3, $r-1 < p-r \le i < p-1$,

$$\begin{aligned}\pi\tilde{h}_i^* &= a_1^2(2\phi^{2i}+\phi^{2(r-1)})\sigma_\pi^4 + a_1a_2(2\phi^{2i+(p-r)}+\phi^{2i-(p-r)}+3\phi^{p+r-2})\sigma_\pi^4 \\ &\quad + a_2^2(\phi^{2i}+2\phi^{2(p-1)})\sigma_\pi^4.\end{aligned}$$

The identity $\sum_{i=j}^{k-1}\phi^{2i} = (1-\phi^{2(k-j)})/(1-\phi^2)$ then provides the sum

$$\begin{aligned}\sum_{i=0}^{p-2}\tilde{\pi}I_{\tilde{g}}\tilde{P}^i\tilde{g} &= \sum_{i=0}^{p-2}\pi\tilde{h}_i^* \qquad (19)\\ &= a_1^2\left\{\left(\frac{1-2\phi^{2(p-1)}}{1-\phi^2}\right)+\phi^{2(r-1)}\left(p+r-\frac{1-2\phi^2}{1-\phi^2}\right)\right\}\sigma_\pi^4 \\ &\quad + a_1a_2\phi^{p-r}\left\{p-r+2\left(\frac{1-\phi^{2(p-1)}}{1-\phi^2}\right)+(3p+r-4)\phi^{2(r-1)}\right\}\sigma_\pi^4 \\ &\quad + a_2^2\left\{\left(\frac{1-\phi^{2(p-1)}}{1-\phi^2}\right)+2(p-1)\phi^{2(p-1)}\right\}\sigma_\pi^4.\end{aligned}$$

To calculate the last term in (12), observe that $\tilde{\pi}I_{\tilde{g}}\tilde{\pi}\tilde{g} = \tilde{\pi}(\tilde{g}^2) = \tilde{\pi}\tilde{h}_0 = \pi\tilde{h}_0^*$, which is given by (18) with $i=0$. Thus,

$$\begin{aligned}\tilde{\pi}I_{\tilde{g}}\{I-\tilde{\Pi}\}\tilde{g} &= \pi\tilde{h}_0^* - (\pi\tilde{g}^*)^2 \qquad (20)\\ &= a_1^2(1+\phi^{2(r-1)})\sigma_\pi^4 + 2a_1a_2\phi^{p-r}(1+\phi^{2(r-1)})\sigma_\pi^4 + a_2^2(1+\phi^{2(p-1)})\sigma_\pi^4.\end{aligned}$$

Putting together (16), (17), (19), and (20), The asymptotic variance (12) is given by

$$\begin{aligned}\sigma_{\tilde{g}}^2 &= a_1^2\left\{\left(\frac{1+\phi^2}{1-\phi^2}\right)+\phi^{2(r-1)}\left(2(r-1)+\frac{1+\phi^2}{1-\phi^2}\right)\right\}\sigma_\pi^4 \qquad (21)\\ &\quad + 2a_1a_2\phi^{p-r}\left\{p-r+\left(\frac{1+\phi^2}{1-\phi^2}\right)+\phi^{2(r-1)}\left(p+r-2+\frac{1+\phi^2}{1-\phi^2}\right)\right\}\sigma_\pi^4 \\ &\quad + a_2^2\left\{\left(\frac{1+\phi^2}{1-\phi^2}\right)+\phi^{2(p-1)}\left(2(p-1)+\frac{1+\phi^2}{1-\phi^2}\right)\right\}\sigma_\pi^4.\end{aligned}$$

Example 3. Under the AR(1) model, a least squares estimator for ϕ arises by regressing one-step-ahead values Φ_{i+1} onto current values Φ_i in the manner of no-intercept simple linear regression. Calculating this from a trajectory of $n+1$ time points, $\Phi_1,\ldots,\Phi_{n+1}$, the estimator is

$$\hat{\phi}_n = \frac{\sum_{i=1}^n \Phi_i\Phi_{i+1}}{\sum_{i=1}^n \Phi_i^2}.$$

Observe that $\hat{\phi}$ is calculated from the partial sums of the previous example with $p=2$ and $r=1$: the partial sums are $S_1 = \sum_{i=1}^n \Phi_i^2$ and $S_2 = \sum_{i=1}^n \Phi_i\Phi_{i+1}$, and the least squares estimator is $\hat{\phi}(S_1,S_2) = S_2/S_1$.

The asymptotic properties of $\hat{\phi}_n$ may therefore be calculated from the asymptotic means, variances, and the covariance of S_1 and S_2 by Taylor expansion. A first-order expansion of $\hat{\phi}(S_1, S_2)$ about (μ_1, μ_2) implies that the asymptotic mean of $\hat{\phi}_n$ is μ_2/μ_1 and its asymptotic variance is

$$nVar[\hat{\phi}] \approx \{\partial_1\hat{\phi}\}^2\sigma_1^2 + 2\{\partial_1\hat{\phi}\}\{\partial_2\hat{\phi}\}\sigma_{12} + \{\partial_2\hat{\phi}\}^2\sigma_2^2. \tag{22}$$

where $\partial_i\hat{\phi} = \partial/(\partial S_i)\hat{\phi}(\mu_1, \mu_2)$. Note that $\partial_1\hat{\phi} = -\mu_2\mu_1^{-2}$ and $\partial_2\hat{\phi} = \mu_1^{-1}$. According to the formulas (17) and (21), these are $\mu_1 = \sigma_\pi^2$ and $\mu_2 = \phi\sigma_\pi^2$,

$$\sigma_1^2 = 2\left(\frac{1+\phi^2}{1-\phi^2}\right)\sigma_\pi^4, \sigma_2^2 = \left\{1 + \phi^2 + 4\left(\frac{\phi^2}{1-\phi^2}\right)\right\}\sigma_\pi^4, \text{ and } \sigma_{12} = 4\left(\frac{\phi}{1-\phi^2}\right)\sigma_\pi^4.$$

The asymptotic mean of $\hat{\phi}_n$ is therefore $\mu_2/\mu_1 = \phi$. Plugging σ_1^2, σ_2^2, and σ_{12} into (22) greatly simplifies the expression and leads to the well known formula for the asymptotic variance of $\hat{\phi}$ given by $nVar[\hat{\phi}] \approx \sigma^2(1-\phi^2)$.

4 Conclusions and discussion

A linear operator framework was established for discrete-time Markov chains and the Drazin inverse was shown to unify objects central to Markov chain theory. This framework was applied to asymptotic variance calculation for univariate and higher-order partial sums, with the formulas expressed in terms of objects governing the original chain, $\{\Phi_i\}$, which presumably would be obtained by statistical analysis.

A potential application is to analysis of a general autoregressive time series. The present study does not go as far as to provide a methodology for the analysis of general autoregressive time series, but our hope is that this technique might be exploited in future applications in order to reduce computational expense. Another extension of these methods would be to continuous-time Markov chains.

We note some existing literature which used asymptotic variance formulas to monitor the convergence of simulated Markov chains to their invariant distributions. For instance, Peskun (1973) studied convergence in chains with finite state space using the fundamental matrix. More recent work by Chauveau and Diebolt (2002) builds upon the formula (10) to study convergence in general state-space chains. We conjecture the present results would be useful in these applications as well.

It is prudent to remark on continuity and approximation of π and $Q^\#$. Continuity results for $Q^\#$ are found in Campbell (1980) and Koliha and Rakočević (1998). These can be extended to the continuity of π through our Proposition 3.*iii*. Meyer (1975) contains an early algorithm for calculating $Q^\#$ for finite state-space Markov chains. Working with abstract Banach spaces, Gonzolez and Koliha (1999) study semi-iterative methods for solving linear equations such as $\pi = \pi P$ in cases where $Q^\#$ exists. More recently, iterative methods for approximating $Q^\#$ and other generalized inverses in general state-space settings are studied in Djordjevic, Stanimirovic, and Wei (2004). We expect the results of these investigations would be helpful in developing efficient computational algorithms for Markov chain applications.

5 Acknowledgements

The authors thank the anonymous referee for his/her helpful suggestions. Dr. Spitzner also thanks Profs Richard Smith and Andrew Nobel for their comments and guidance.

References

[1] Brockwell, P. J., and Davis, R. A. (1991), *Time Series: Theory and Methods (second edition),* New York: Springer. MR1093459

[2] Campbell, S. L. (1980), Continuity of the Drazin inverse, *Linear and Multilinear Algebra,* 8:265-268. MR0560571

[3] Campbell, S. L., and Meyer, C. D. (1979), *Generalized Inverses of Linear Transformations,* London: Pitman. MR0533666

[4] Chauveau, D., Diebolt, J. (2003), Estimation of the asymptotic variance in the CLT for Markov chains, *Stochastic Models,* 19(4): 449-465. MR2013368

[5] Chan, K. S., and Geyer, C. J. (1994), Discussion of"Markov chains for exploring posterior distributions," *Annals of Statistics,* 22: 1747-1758. MR1329166

[6] Cogburn, R. (1972), The central limit theorem for Markov processes, *Proceedings of the 6th Berkeley symposium on mathematical statistics and probability,* 485-512.

[7] Djordjevic, D. S., Stanimirovic, P. S., and Wei, Y. (2004), The representation and approximations of outer generalized inverses, *Acta Math. Hungar.,* 104(1-2): 1-26. MR2069959

[8] François, O. (1999), On the spectral gap of a time reversible Markov chain, *Probability in Engineering and Informational Sciences,* 13:95-101. MR1666392

[9] Geyer, C. J. (1992), Practical Markov chain Monte Carlo (with discussion), *Statistical Science,* 7:473-511.

[10] Gonzalez, C. N., and Koliha, J. J. (1999), Semi-iterative methods for the Drazin inverse solution of linear equations in Banach spaces, *Numer. Funct. Anal. Optim.,* 20(5-6): 405-418. MR1704952

[11] Glynn, P. W., and Meyn, S. P. (1996), A Liapunov Bound for Solutions of the Poisson Equation, *Annals of Probability,* 24:916-931. MR1404536

[12] Hobart, G. L., Jones, G. L., Presnell, B., and Rosenthal, J. S. (2002), On the applicability of regenerative simulation in Markov chain Monte Carlo, *Biometrika,* 89:731-743. MR1946508

[13] Jones, G. L. (2004), On the Markov chain central limit theorem, *Probability Surveys,* 1:299-320. MR2068475

[14] Kemeny, J. G., and Snell, J. L. (1960), *Finite Markov Chains*, New York: Van Nostrand. MR0115196

[15] Kipnis, C., and Varadhan, S. R. S. (1986), Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions, *Communications in Mathematical Physics,* 104:1-19. MR0834478

[16] Koliha, J. J. (1996a), A generalized Drazin inverse, *Glasgow Mathematics Journal,* 38:367-381. MR1417366

[17] Koliha, J. J. (1996b), Isolated spectral points, *Proceedings of the American Mathematical Society,* 124(11):3417-3424. MR1342031

[18] Koliha, J. J., and Rakočević, V. (1998), Continuity of the Drazin inverse II, *Studia Mathematica,* 131(2):167-177. MR1636348

[19] Lawler, G. F., and Sokal, A. D. (1988), Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality, *Transactions of the American Mathematical Society,* 309(2):557-580. MR0930082

[20] León, C. A. (2001), Maximum asymptotic variance of sums of finite Markov chains, *Statistics and Probability Letters,* 54:413-415. MR1861387

[21] Meyer, C. D. (1975), The role of the group generalized inverse in the theory of finite Markov chains, *SIAM Review,* 17(3): 443-464. MR0383538

[22] Meyn, S. P., and Tweedie, R. L. (1993), *Markov Chains and Stochastic Stability,* London, U.K.: Springer-Verlag. MR1287609

[23] Nummelin, E. (1984), *General Irreducible Markov Chains and Non-negative Operators,* New York: Cambridge University Press. MR0776608

[24] Nummelin, E. (2002), MC's for MCMC'ists, *International Statistical Review,* 70:215-240.

[25] Peskun, P. H. (1973), Optimum Monte-Carlo sampling using Markov chains, *Biometrika,* 60(3):607-612. MR0362823

[26] Robert, C. P. (1995), Convergence control methods for Markov chain Monte Carlo algorithms, *Statistical Science,* 10:231-253. MR1390517

[27] Robert, C. P., and Casella, G. (1999), *Monte Carlo Statistical Methods,* New York: Springer. MR1707311

[28] Roberts, G. O., and Tweedie, R. L. (1996), Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms, *Biometrika,* 83:95-110. MR1399158

[29] Roberts, G. O., and Rosenthal, J. S. (2004), General state space Markov Chains and MCMC algorithms, *Probability Surveys,* 1:20-71. MR2095565

[30] Tierney, L. (1994), Markov chains for exploring posterior distributions (with discussion), *Annals of Statistics,* 22:1701-1762. MR1329166

[31] Tóth, B. (1986), Persistent random walks in random environment. *Probability Theory and Related Fields,* 71:615-625. MR0833271