Mathematical Methods in Operations Research 2005
Special issue in honor of *Arie Hordijk*

Ave, Lyapunov functions!

Flora Spieksma

University of Leiden

Received: date / Revised version: date

Abstract Contractive Lyapunov functions play a non neglectable role in convergence properties of Markov chains. We will discuss properties of the existence of these functions, as well as their impact on some problems.

1 Introduction

My cooperation with Arie Hordijk has been spiced by Lyapunov function methods. The first application of such functions has been in the context of Markov decision chains, with applications to the optimal control of queues. My main goal for the PhD work was to prove an intriguing conjecture from Rommert Dekker's thesis, my predecessor PhD student with Arie, which is largely a version of Theorem 1 for Markov decision chains.

Both Arie and myself have continued working on various applications of "contractive" Lyapunov functions till this day. I will give an overview of some of our issues of interest. Contractive Lyapunov functions form the essential basis for all this work, although this is not always directly discernible. For this reason, I have taken the existence of this type of function as the starting point and I have thrown in an occasional generalisation that simply fits the conditions required.

2 Model and main Lyapunov function criterion

Our basic model is a discrete time Markov chain $\{\xi_t\}_t$ on the countable state space $\boldsymbol{S}$ with stationary transition probabilities $p_{xy} = \mathsf{P}_x(\xi_{t+1} = y | \xi_t = x)$. We assume the Markov chain to be irreducible and aperiodic. P_x, E_x stand for conditional probability and conditional expectation given $\xi_0 = x$, and $\xi_t(x)$ for the state of the chain at time t given $\xi_0 = x$.

Our basic criterion is the following. The non-negative function $f : \boldsymbol{S} \to \mathbf{R}_+$ is called a **(contractive) Lyapunov function**, if there exist an *exception set* $A \subset \boldsymbol{S}$, with $A, A^c := \boldsymbol{S} \setminus A \neq \emptyset$; a bounded step function $k : \boldsymbol{S} \to \mathbf{Z}_+$, i.e., $\sup_x k(x) < \infty$, and constants $\gamma, c \geq 0$, such that

$$\mathsf{E}_x\{f(\xi_{k(x)})\} \leq \exp\{-\gamma\} f(x), \quad x \notin A; \tag{1}$$

$\mathsf{E}_x\{f(\xi_{k(x)})\} < \infty$ for $x \in A$, and $\mathsf{E}_x\{f(\xi_1)\} \leq cf(x)$, $x \notin A$.

These are rather strong conditions that roughly require the one step jumps to have uniformly (in the states of A^c) convergent Laplace Stieltjes transforms in a neighbourhood of 0. However, in the spirit of Arie's work, these conditions do allow considering the restricted transition matrix as a bounded linear operator on a normed Banach space into itself.

Let ${}^AP = (p_{xy})_{x,y\notin A}$ denote the transition matrix restricted to A^c. Introduce the f-norm $\|g\|_f = \sup_x |g(x)|/f(x)$ and denote by $\mathsf{L}^\infty(A^c, f) = \{g : A^c \to \mathbf{R} \,|\, \|g\|_f < \infty\}$ the space of functions on A^c with finite f-norm.

The conditions imply that ${}^AP$ acts as a bounded linear operator from $\mathsf{L}^\infty(A^c, f)$ to itself with norm $\|{}^AP\|_f = \sup_{x\notin A} \sum_{y\notin A} p_{xy} f(y)/f(x)$. In particular, one can show (see [11]) that $\|{}^AP^{(t)}\|_f \leq \tilde{c} \exp\{-\gamma t\}$ for some constant $\tilde{c}$ that can be determined from the above given constants. Here ${}^AP^{(t)} = \left({}^AP\right)^t$ stands for the t-step reduced transition matrix.

Let $\tau_A = \min\{t > 0 \,|\, \xi_t \in A\}$. The following lemma follows directly.

Lemma 1 *If there exists $x \in A^c$ with $f(x) < \inf_{a\in A} f(a)$, then $\mathsf{P}_x\{\tau_A < \infty\} < 1$, implying in turn that the Markov chain is transient. If $\inf_{x\notin A} f(x) > 0$, then $\tau_A < \infty$ with probability 1 and $\mathsf{E}\tau_A < \infty$, for* all *initial states $\xi_0 = x \notin A$.*

3 Exponential convergence properties

Now if we assume *finiteness* of the exception set A, then we can pull back exponential convergence (to 0) of the reduced transition matrix to exponential convergence of the un-restricted matrix, in the following sense. The proof uses first entrance-last exit decomposition to the set A.

Theorem 1 *[7], [11]. Let $f : \boldsymbol{S} \to \mathbf{R}_+$. The following two conditions are equivalent.*

i) The function f is a contractive Lyapunov function for a finite *exception set $A \subset \boldsymbol{S}$.*

ii) There exist positive constants β, b, such that $P : \mathsf{L}^\infty(\boldsymbol{S}, f) \to \mathsf{L}^\infty(\boldsymbol{S}, f)$ is an f-bounded operator and

$$\|P^{(t)} - \mathsf{L}\|_f \leq b \cdot \exp\{-\beta t\},$$

where $\mathsf{L} = \lim_{t\to\infty} P^{(t)}$, which is well-known (see [2]) to exist.

If $\inf_x f(x) = 0$ *then* $\mathsf{L} \equiv 0$ *and so the Markov chain is exponentially transient; if* $\inf_x f(x) > 0$ *then the Markov chain is exponentially ergodic. Reversely, if the Markov chain is exponentially ergodic or transient, then a contractive Lyapunov function with the exception set* A *finite, exists.*

A number of consequences of the above theorem can be found in the above references. These include reward chains as special cases of Markov decision chains.

More relevant are the following largely open problems: (i) how to compute β, given a contractive Lyapunov function with finite exception set; (ii) given the exponential convergence rates of $p_{xy}^{(t)}$, is there a normed space, with exponential convergence rate related to these individual convergence rates? Limited results have been provided in [10]. Lund and Tweedie[8] have provided the sharpest answer so far: for so-called *stochastically monotone chains* the answer to (ii) is yes and any $\beta < \gamma$ suffices, with γ from equation 1, provided $k(x) \equiv 1$ and A consists of one minimal state.

4 Transient case: Almost closed set structure

A good reason for allowing the exception set A to be infinite, is when studying the different directions into which a transient Markov chain may escape to infinity. This problem plays a role in studying the Poisson or the Martin boundary of the chain. It also comes up when studying fluid limits (see [3], [5]). Arie and his PhD student Nikolay Popov found this problem to emerge as a part of methods for explicitly computing Large Deviation estimates in e.g., [6].

To describe the concept of escape direction more precisely, we have to introduce Blackwell's concept ([1]) of an *almost closed set.* The set $A \subset \boldsymbol{S}$ is called a.c. (almost closed), whenever for some state x (and hence for all)

$$\mathsf{P}_x\big(\limsup_{t\to\infty}\{\xi_t \in A\}\big) = \mathsf{P}_x\big(\liminf_{t\to\infty}\{\xi_t \in A\}\big)(= \lim_{t\to\infty} \mathsf{P}_x(\xi_t \in A)) > 0.$$

It is called *transient*, whenever $\mathsf{P}_x\big\{\limsup_{t\to\infty}\{\xi_t \in A\}\big\} = 0$. An a.c. set A is called *atomic* if it does not contain two disjoint a.c. subsets. It is called *completely non-atomic* if it does not contain any atomic a.c. subsets. Now Blackwell proved that there is a countable partition $\{A_0, A_1, \ldots\}$ of the $\boldsymbol{S}$, such that $A_1, \ldots$ are atomic a.c. sets (if present) and A_0 (if present) is completely non-atomic. This decomposition is unique upto transient sets. Moreover,

$$\begin{aligned} 1 &= \sum_{n=0}^{\infty} \mathsf{P}_x\big(\limsup_{t\to\infty}\{\xi_t \in A_n\}\big) \\ &= \sum_{n=0}^{\infty} \mathsf{P}_x\big(\liminf_{t\to\infty}\{\xi_t \in A_n\}\big) = \sum_{n=0}^{\infty} \lim_{t\to\infty} \mathsf{P}_x(\xi_t \in A_n), \end{aligned} \tag{2}$$

the latter expression being the crucial one that we often need. Now one can extend Lemma 1 as follows (see [5]).

Lemma 2 *Let f be a contractive Lyapunov function with exception set A.*

i) Let there be a state $x \in A^c$, with $f(x) < \alpha = \inf_{a\in A} f(a)$. Then the set $\{y|f(y) < \alpha - \epsilon\}$ is a.c. for any $0 \le \epsilon < \alpha$.

ii) Suppose that $\inf_{x\in A^c} f(x) > 0$, and that $\inf_{x\in A} \mathsf{P}_x\big(\liminf_{t\to\infty}\{\xi_t \in A\}\big) > 0$. Then the set A^c is transient.

In the second case, the Markov chain need *not* be transient!

Often f is a contractive Lyapunov function with a finite exception set, such that $\sup_x f(x) = \infty$, and $\inf_x f(x) = 0$. Then the above lemma allows to limit the state space to the smaller a.c. part $\{x|f(x) < \alpha\}$ of it, for some $\alpha > 0$.

Now we have not yet addressed the problem of constructing a contractive Lyapunov function. We wish to identify atomic a.s. subsets and to conduct a more refined study of where the process actually moves in the a.c. set. For instance, the latter may stem from a study of fluid limits.

Write $U(x;t)$ for the cumulative sum of expected jumps along the paths of the Markov chain:

$$U(x;t) = x + \sum_{n=1}^{t} \mathsf{E}_x(\xi_n - \xi_{n-1}|\xi_{n-1}), \quad t = 0, 1, \dots$$

Clearly, $\xi_t(x) - U(x;t)$ is a martingale. The conditions of Theorem 2 then imply that $\big(\xi_t(x) - U(x;t)\big)/T \to 0$ a.s. (see [4]). As a consequence, for studying the behaviour of $\xi_{[tN]}([xN])/N$, $N \to \infty$, we may as well study $U([xN];[tN])/N$, $N \to \infty$. Very often the jump distributions $\{\,p_{xy}\}_{y\in\boldsymbol{S}}$ are a finite collection and so the latter limit is more easily studied and very often can be shown to exist in distribution. Say it has a limit random vector $u(x;t)$, which will be called the *fluid paths*.

Assume that $\boldsymbol{S} \subset \mathbf{Z}^p$. By $\mathsf{conv}(A)$ we mean the convex hull of $A \subset \boldsymbol{S}$ in $\mathbf{R}^p$. The following conjecture has been proved in [5] for some low-dimensional queueing examples. The relevance of equation 2 emerges here.

Conjecture Suppose that there are no non-atomic a.c. sets. Then to each a.c. set C corresponds precisely one 'path' $u(0;t)$, $t \ge 0$, such that $u(0;t) \subset \mathsf{conv}(C)$ for $t \ge T$, for some finite time T. Moreover,

$$\mathsf{P}_x\Big(\lim_{N\to\infty} \frac{\xi_{[tN]}(x)}{N} = u(0;t)\Big) = \lim_{N\to\infty} \mathsf{P}_x(\xi_N \in C).$$

Now these fluid paths can be used for constructing suitable contractive Lyapunov functions, through sub-additive ones.

5 Sub-additive Lyapunov functions

There is the very simple trick of transforming a sub-additive relation into a contractive one by taking exponentials. This was put to work in the context of Lyapunov functions in [12], based on ideas from [9]. The following result can be shown by using a Taylor expansion.

Theorem 2 *Let g be a sub-additive Lyapunov function, that is, there exists an exception set $A \subset \boldsymbol{S}$, $A, A^c \neq \emptyset$, a bounded step function $k : \boldsymbol{S} \to \mathbf{Z}_+$, and a constant $\epsilon > 0$, such that*

$$\mathsf{E}_x\big(g(\xi_{k(x)})\big) \leq g(x) - \epsilon, \quad x \notin A, \tag{3}$$

and $\mathsf{E}_x\big(g(\xi_{k(x)})\big) < \infty$, $x \in A$. Suppose that further that $\boldsymbol{S}$ is endowed with a norm $\|\cdot\|$, such that g is Lipschitz w.r.t this norm: $|g(x)-g(y)| \leq c\cdot\|x-y\|$ for some constant c. Moreover, $\sup_x \sum_y p_{xy} \exp\{\delta\|y-x\|\} < \infty$, for some constant $\delta > 0$. Then $f(x) = \exp\{hg(x)\}$ is a contractive Lyapunov function for all sufficiently small $h > 0$.

The advantage is that sub-additive Lyapunov functions are easier to construct (cf. [12]) Allowing the step function k to be general instead of $k(\cdot) \equiv 1$, has a smoothing effect on the Lyapunov function to be constructed.

For a positive recurrent chain, a minimal sub-additive Lyapunov function is $g(x) = \epsilon \mathsf{E}_x \tau_A$. This interpretation of 'expected hitting time' is important. Knowing the fluid paths, one can take the expected time to reach a set A along the fluid path or along the time-reversed fluid path as a basis for this construction.

References

1. D. Blackwell (1942), On idempotent Markoff chains. *Ann. Mathematics* **43**, 560–567.
2. K.L. Chung (1960), *Markov Chains with Stationary Transition Probabilities.* Springer-Verlag, Berlin.
3. G. Fayolle, V.A. Malyshev, and M.V. Menshikov (1995), *Constructive Theory of countable Markov Chains.* Cambridge University Press, Cambridge.
4. P. Hall and C.C. Heyde (1980), *Martingale limit Theory and its Applications.* Academic Press, San Diego.
5. A. Hordijk, N. Popov, and F.M. Spieksma (2002), Discrete scattering and some simple non-simple face-homogeneous random walks. Technical report MI 2002-26, University of Leiden. submitted for publication in Appl. Prob.
6. A. Hordijk and N.V. Popov (2003), Large deviation bounds for face-homogeneous random walks in the quarter plane. *PEIS* **173**, 397–406.
7. A. Hordijk and F.M. Spieksma (1992), On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Adv. Appl. Prob.* **24**, 343–376.
8. R.B. Lund and R.L. Tweedie (1996), Geometric convergence rates for stochastically ordered Markov chains. *Maths of Operat. Res.* **21**, 182–194.
9. V.A. Malyshev and M.V. Menshikov (1981), Ergodicity, continuity and analyticity of countable Markov chains. *Trans. Moscow Math. Soc.* **1**, 1–48.
10. F.M. Spieksma (1992), Spectral conditions and bounds for the rate of convergence of countable Markov chains. Technical report No. TW-92-11, Dept. of Mathematics and Comp. Science, Leiden Univ.
11. F.M. Spieksma, Lyapunov functions for Markov chains with applications to face-homogeneous random walks. *Internal communication.*
12. F.M. Spieksma and R.L. Tweedie (1994), Strengthening ergodicity to geometric ergodicity of Markov chains. *Stoch. Models* **10**, 45–75.