LMS/EPSRC Short Course Stability, Coupling Methods and Rare Events

Heriot-Watt University, Edinburgh, 4-9 September 2006

Lecture Notes on:

Lyapunov function methods

Serguei Foss and Takis Konstantopoulos

1 History

Originally, Lyapunov functions were developed by A. Lyapunov in 1899 for the study of stability of dynamical systems described by ODE's, mainly motivated by mechanical systems. Since then, the methods based on Lyapunov functions have been extended to study stability of dynamical systems of all kinds (chaotic systems, control systems, stochastic systems, discrete systems, etc.)

2 Introduction

Our goal in this part of the lectures is to exemplify why and how Lyapunov functions work for (Harris) Markov processes in a general state space. To avoid technicalities (which have not fully resolved) we shall avoid continuous time.

The field of applications that interests us is those stochastic systems that appear to have "simple evolution" "away from boundaries". Highly-nonlinear but smooth stochastic systems are also interesting but more well-understood than our cases.

We start with a few examples of Lyapunov functions, for some classical dynamical systems.

Linear systems A linear system in $\mathbb{R}^d$ is another name for the ODE

$$\dot{x} = Ax$$

where A is a $d \times d$ matrix. The system is asymptotically stable (definition) if for all x_0 in a neighborhood of the origin the trajectory starting from x_0 tends to 0 as $t \to \infty$. This holds if and only if all eigenvalues have real part strictly smaller than 0. Consider the scalar function $V(x) = x^\mathsf{T} Q x$, $x \in \mathbb{R}^d$, where Q is a positive definite matrix. Think of this as an "energy function". Since energy of a dissipative system cannot increase, we expect that $V(x)$ should decrease along a trajectory. To see if this is true, we differentiate along a trajectory and obtain

$$\frac{d}{dt} V(x) = x^\mathsf{T} A^\mathsf{T} Q x + x^\mathsf{T} Q A x.$$

We want this to be (strictly) negative. But this is also a quadratic form. So, equivalently, we want to have find a positive definite matrix P such that

$$A^\mathsf{T} Q + Q A^\mathsf{T} + P = 0.$$

It is well known that the latter (linear) equation has a solution in P, Q, s.t. both P and Q are positive definite if and only if the eigenvalues of A have real part strictly less than 0.

Asymptotic stability of nonlinear systems Consider the autonomous nonlinear system described by an ODE

$$\dot{x} = f(x), \quad x \in W \text{ open } \subseteq \mathbb{R}^d,$$

where f is Lipschitz continuous on W. Suppose that 0 is a stationary point: $f(0) = 0$. We say that the system is asymptotically stable at 0, whenever there is a neighbourhood of zero such that if we start at an x_0 within this neghboourhood, the trajectory tends to 0 as $t \to \infty$.

Theorem 1 (Lyapunov's theorem)**.** *If there exists a differentiable function $V : W_1 \to \mathbb{R}$, where W_1 an open neighborhood of 0, $W_1 \subseteq W$, such that $V(0) = 0$, $V(x) > 0$ for all $x \in W_1 \setminus \{0\}$, and $\langle V'(x), f(x)\rangle < 0$ for all $x \in W_1 \setminus \{0\}$ then asymptotic stability at 0 is guaranteed.*

The reason for this is that $V(x)$ acts like a norm which is decreasing along a trajectory at a rate bounded away from zero.

Example: nonlinear (van der Pol) oscillator This is an oscillator with nonlinear damping:

$$\ddot{x} + x + \varepsilon(\dot{x} - b\dot{x}^3) = 0,$$

where $\varepsilon, b > 0$. Let $x_1 = x, x_2 = \dot{x}$, and then

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 - \varepsilon(x_2 - bx_2^3) \end{pmatrix}$$

Define

$$V(x_1, x_2) = \frac{1}{2}\left(x_1^2 + x_2^2\right).$$

Then

$$V' = \left(\frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_1}\right) = (x_1, x_2),$$

$$\langle V', f\rangle = x_1x_2 - x_1x_2 + \varepsilon(bx_2^4 - x_2^2) = -\varepsilon(x_2^2 - bx_2^4)$$

Hence if $x_2 < 1/\sqrt{b}$, then 0 is asymptotically stable. (The case $b = 0$ corresponds to a linear oscillator with damping which is globally asymptotically stable without any constraint on the velocity.)

Recursions We can ask about the stability of a recursion of the form

$$x_{n+1} = f(x_n), \quad x_n \in \mathbb{R}^d.$$

Lyapunov function methods can be defined here as well.

Non-asymptotic stability of general dynamical systems Asymptotic stability is often too restrictive. For instance, the periodic elliptic motion of a mass around a fixed center is the result of a stable but not asymtotically so system. Or, if we have a differential equation of the form $\dot{x} = f(x,t)$, with explicit dependence on time, it is more reasonable to talk about non-asymptotic stability. Imagine that $(\theta_{s,t}, 0 \le s \le t < \infty)$ is a flow in some metric space M. This means that for all $s < t$, the map $\theta_{s,t} : M \to M$ is measurable and that

$$\theta_{s,u} = \theta_{s,t} \circ \theta_{t,u} \quad \theta_{t,t} = \mathrm{id}_M.$$

Suppose that $\mathbf{0} \in M$ is a stationary point, namely, $\theta_{s,t}\mathbf{0} = \mathbf{0}$ for all (initial times s) and all $t \ge s$. We can say that $\mathbf{0}$ is stable if for each open neighbourhood U of $\mathbf{0}$ there is another open neghbourhood V of $\mathbf{0}$ such that, for all s and all $x \in V$ we have $\theta_{s,t}x \in U$ for all $t \ge s$.

3 Lyapunov functions for Markov processes

The problem of stability for a stochastic system described by a Markov process often (most of the time?) boils down to proving that some set is positive recurrent.

The setup Our object of study is a Markov process (X_n) with values in some general state space S which will be assumed to be Polish.[1]

A time-homogeneous Markov process with values in S can be described either by its transition probability kernel

$$P(x, B) = P(X_{n+1} \in B \mid X_n = x)$$

or by a stochastic recursion

$$X_{n+1} = f(X_n, \xi_n),$$

where the ξ_n are i.i.d. random variables. The existence of this explicit representation is simple when S is countable (and technical when S is a Polish space).

Stationarity This means existence of a stationary Markov process with the given transition kernel or, equivalently, existence of a stationary probability measure π on S:

$$\pi(B) = \int_S \pi(dx) P(x, B).$$

Uniqueness is also often desirable.

Stability This means convergence of X_n, as $n \to \infty$, in some sense.

We first realise that a.s. convergence is impossible, except in trivial situations when we have absorbing states.[2]

[1] Polish=a separable topological space which is complete under some metric.

[2] Compare, however, with time-inhomogeneous chains, such as the ones encountered in simulated annealing, where a.s. convergence IS possible AND desirable.

The weakest notion of converngence is that, for some x, the probability $P_x(X_n \in \cdot)$ converges (weakly), as $n \to \infty$, to some honest probability.

A stonger notion is to claim the above for all $x \in S$.

An even stronger notion is convergence in total variation.

Drift Let $V : S \to \mathbb{R}$ be some function on the state space S of a Markov process (X_n). The drift of V in n steps is defined by

$$DV(x, n) := E_x[V(X_n) - V(X_0)]$$

(provided the expectations exist)[3] There is a space and a time argument in $DV(x, n)$. It is much more general and convenient to define the drift for a state-dependent time-horizon, i.e. make n a function of x.

So, given a function $g : S \to \mathbb{N}$, we let

$$DV(x, g) := E_x[V(X_{g(x)}) - V(x)].$$

(Positive) Recurrence of a set For a measurable set $B \subseteq S$ define

$$\tau_B = \inf\{n \geq 1 : \ X_n \in B\}$$

to be the first return time[4] to B.

- The set B is called *recurrent* if

$$P_x(\tau_B < \infty) = 1, \quad \text{for all } x \in B.$$

- It is called *positive recurrent* if

$$\sup_{x \in B} E_x \tau_B < \infty.$$

It is this last property that is determined by a suitably designed Lyapunov function.

Roughly speaking, we want to prove that if we can find a function V (the Lyapunov function) such that the drift is negative outside a set, then the set is positive recurrent.

This is the content of Theorem 2 below. That this property can be translated into a stability statement is the subject of a next lecture.

First we impose a number of assumptions.

[3] In continuous time, for example for a Markov jump process (X_t), the drift of V in t time units equals $DV(x, t) = E_x[V(X_t) - V(X_0)]$. This can also be written as $(P_t V)(x)$ in operator notation. Under additional assumptions, $DV(x, t) = \int_0^t QV(X_s)ds$, where Q is the generator of the process. So the (infinitesimal) drift of V is simply $QV(x)$, and if $q(x, y)$ denotes the rate from state x to state y, we have $QV(x) = \sum_y [V(y) - V(x)]q(x, y)$. Note that, if V is smooth enough, we can write $E_x[V(X_t) - V(X_0)] = E_x \int_0^t \frac{d}{ds} V(X_s)ds$, from which it is, at least intuitively, clear why $QV(x)$ corresponds to differentiation in time and why this notion of drift is the correct stochastic analogue of what we defined earlier for deterministic ODEs. In discrete time we avoid the "technicalities" of taking derivatives or of defining the process from its generator, so the whole exposition here will be in discrete time. It should be noticed, however, that in applications, Markov processes are often more natural in continuous rather than in discrete time.

[4] This τ_B is a random variable. Were we working in continuous time, this would not, in general, be true, unless the paths of X and the set B were sufficiently "nice" (another instance of what technical complexities may arise in a continuous-time setup).

Assumptions

(L0) V is unbounded from above: $\sup_{x \in S} V(x) = \infty$.

(L1) h is bounded from below: $\inf_{x \in S} h(x) > -\infty$.

(L2) h is eventually positive: $\underline{\lim}_{V(x) \to \infty} h(x) > 0$.

(L3) g is locally bounded from above: $\sup_{V(x) \le N} g(x) < \infty, \quad$ for all $N > 0$.

(L4) g is eventually bounded by h: $\overline{\lim}_{V(x) \to \infty} g(x)/h(x) < \infty$.

Theorem 2. *Suppose that the drift of V in $g(x)$ steps satisfies*

$$E_x[V(X_{g(x)}) - V(X_0)] \le -h(x),$$

where V, g, h satisfy (L0)–(L4). *Let*

$$\tau \equiv \tau_N = \inf\{n \ge 1 : \ V(X_n) \le N\}.$$

Then there exists $N_0 > 0$, such that for all $N > N_0$ and any $x \in S$, we have

$$E_x \tau < \infty,$$
$$\sup_{V(x) \le N} E_x \tau < \infty.$$

Proof. We follow an idea that is essentially due to Tweedie (1976). From the drift condition, we obviously have that $V(x) - h(x) \ge 0$ for all x. We choose N_0 such that $\inf_{V(x) > N_0} h(x) > 0$. Then, for, $N \ge N_0$, we set

$$d = \sup_{V(x) > N} g(x)/h(x), \quad -H = \inf_{x \in S} h(x), \quad c = \inf_{V(x) > N} h(x).$$

We define an increasing sequence t_n of stopping times recursively by

$$t_0 = 0, \quad t_n = t_{n-1} + g(X_{t_{n-1}}), \quad n \ge 1.$$

By the strong Markov property, the variables

$$Y_n = X_{t_n}$$

form a (possibly time-inhomogeneous) Markov chain with, as easily proved by induction on n, $E_x V(Y_{n+1}) \le E_x V(Y_n) + H$, and so $E_x V(Y_n) < \infty$ for all n and x. Define the stopping time

$$\gamma = \inf\{n \ge 1 : \ V(Y_n) \le N\} \le \infty,$$

for which

$$\tau \le t_\gamma, \text{ a.s.},$$

and so proving $E_x t_\gamma < \infty$ is enough. Let $\mathscr{F}_n$ be the sigma field generated by $Y_0, \ldots, Y_n$. Note that γ is a "predictable" stopping time in that $\mathbf{1}(\gamma \ge i) \in \mathscr{F}_{i-1}$ for all i. We define the "cumulative energy" between 0 and $\gamma \wedge n$ by

$$\mathscr{E}_n = \sum_{i=0}^{\gamma \wedge n} V(Y_i) = \sum_{i=0}^{n} V(Y_i) \mathbf{1}(\gamma \ge i),$$

and estimate the change $E_x(\mathscr{E}_n - \mathscr{E}_0)$ (which is finite) in a "martingale fashion":[5]

$$\begin{aligned}
E_x(\mathscr{E}_n - \mathscr{E}_0) &= E_x \sum_{i=1}^{n} E_x(V(Y_i)\mathbf{1}(\gamma \ge i) \mid \mathscr{F}_{i-1}) \\
&= E_x \sum_{i=1}^{n} \mathbf{1}(\gamma \ge i) E_x(V(Y_i) \mid \mathscr{F}_{i-1}) \\
&\le E_x \sum_{i=1}^{n} \mathbf{1}(\gamma \ge i) E_x(V(Y_{i-1}) - h(Y_{i-1}) \mid \mathscr{F}_{i-1}) \\
&\le E_x \sum_{i=1}^{n+1} \mathbf{1}(\gamma \ge i-1) E_x(V(Y_{i-1}) - h(Y_{i-1}) \mid \mathscr{F}_{i-1}) \\
&= E_x \mathscr{E}_n - E_x \sum_{i=0}^{n} h(Y_i)\mathbf{1}(\gamma \ge i),
\end{aligned}$$

where we used that $V(x) - h(x) \ge 0$ and, for the last inequality, we also used $\mathbf{1}(\gamma \ge i) \le \mathbf{1}(\gamma \ge i-1)$ and replaced n by $n+1$. From this we obtain

$$E_x \sum_{i=0}^{n} h(Y_i)\mathbf{1}(\gamma \ge i) \le E_x V(X_0) = V(x). \tag{1}$$

Assume $V(x) > N$. Then $V(Y_i) > N$ for $i < \gamma$, by the definition of γ, and so

$$h(Y_i) \ge c > 0, \quad \text{for } i < \gamma, \tag{2}$$

by the definition of c. Also,

$$h(Y_\gamma) \ge -H, \tag{3}$$

by the definition of H. Write $h(Y_i) = h(Y_i)^+ - h(Y_i)^-$ and use (2) and (3) in (1) to obtain

$$E_x \sum_{i=0}^{n} h(Y_i)\mathbf{1}(\gamma > i) \le V(x) + H.$$

Using the monotone convergence theorem and (2), we have

$$cE_x\gamma \le V(x) + H < \infty.$$

Using $h(x) \ge dg(x)$ for $V(x) > N$, we also have

$$\sum_{i=0}^{\gamma-1} h(Y_i) \ge d \sum_{i=0}^{\gamma-1} g(Y_i) = dt_\gamma,$$

whence $t_\gamma < \infty$, a.s., and so

$$E_x\tau \le E_x t_\gamma \le \frac{V(x) + H}{d}.$$

It remains to see what happens if $V(x) \le N$. By conditioning on Y_1, we have

$$E_x\tau \le V(x) + E_x(d^{-1}(V(Y_1) + H)\mathbf{1}(V(Y_1) > N)) \le V(x) + d^{-1}H + d^{-1}(V(x) + H).$$

[5] albeit we do not make use of explicit martingale theorems

Hence,

$$\sup_{V(x)\le N} E_x\tau \le N + d^{-1}(2H+N).$$

□

Discussion: The theorem we just proved shows something quite strong about the set $B_N = \{x \in S:\ V(x) \le N\}$. Namely, this set is *positive recurrent*. It is worth seeing that the theorem is a generalization of many more standard methods.

I. Pakes's lemma: This is the case above with $S = \mathbb{Z}$, $g(x) = 1$ and $h(x) = \varepsilon - C_1\mathbf{1}(V(x) \le C_2)$.

II. The Foster-Lyapunov criterion: Here S is general, and $g(x) = 1$ and $h(x) = \varepsilon - C_1\mathbf{1}(V(x) \le C_2)$, Equivalently, the Foster-Lyapunov criterion seeks a function V such that

$$E_x(V(X_1) - V(X_0)) \le -\varepsilon < 0, \qquad \text{when } V(x) > C_2,$$

and

$$\sup_{V(x)\le C_2} E_x V(X_1) < \infty.$$

III. Dai's criterion: When $g(x) = \lceil V(x)\rceil$ (where $\lceil t\rceil = \inf\{n \in \mathbb{N}:\ t \le n\}$, $t > 0$), and $h(x) = \varepsilon V(x) - C_1\mathbf{1}(V(x) \le C_2)$, we have Dai's criterion which is the same as the "fluid limits" criterion. More on this will be seen at a next lecture.

IV. The Meyn-Tweedie criterion: Finally, when $h(x) = g(x) - C_1\mathbf{1}(V(x) \le C_2)$ we have the Meyn-Tweedie criterion.

V. Fayolle-Malyshev-Menshikov: Similar state-dependent drift conditions, for countable Markov chains, were considered by these authors.

The indispensability of the "technical" conditions. It is clear why (L0)–(L3) are needed. As for condition (L4), this is not only a technical condition. Its indispensability can be seen in the following simple example: Consider $S = \mathbb{N}$, and transition probabilities

$$p_{1,1} = 1, \quad p_{k,k+1} \equiv p_k, \quad p_{k,1} = 1 - p_k \equiv q_k, \quad k = 2, 3, \ldots,$$

where $0 < p_k < 1$ for all $k \ge 2$ and $p_k \to 1$, as $k \to \infty$. Thus, jumps are either of size $+1$ or $-k$, till the first time state 1 is hit. Assume

$$q_k = 1/k, \quad k \ge 2.$$

Then 1 is an absorbing state, and there is $C > 0$, such that

$$P(X_{n+1} = X_n + 1 \text{ for all } n) \le C\exp\left(-\sum_k q_k\right) = 0.$$

But, for $\tau = \inf\{n:\ X_n = 1\}$,

$$\sum_n P(\tau \ge n) \ge \sum_n \exp\sum_2^n q_i \sim \sum_n \frac{1}{n} = \infty.$$

Therefore, the Markov chain cannot be positive recurrent. Take now

$$V(k) = \log(1 \vee \log k), \quad g(k) = k^2.$$

We can estimate the drift and find

$$E_k[V(X_{g(k)}) - V(k)] \le -h(k), \tag{4}$$

where $h(k) = c_1 V(k) - c_2$, and c_1, c_2 are positive constants. It is easily seen that (L1)-(L3) hold, but (L4) fails. This makes Theorem 2 inapplicable in spite of the negative drift (4). Physically, the time horizon $g(k)$ over which the drift was computed is far too large compared to the estimate $h(k)$ for the size of the drift itself.

4 Instability criteria

We first give a simple instability criterion due to Tweedie.

Theorem 3. *Suppose there is a non-constant function* $V : S \to \mathbb{R}_+$ *such that*
$DV(x, 1) \ge 0$ *when* $V(x) \ge K$, *for some* $K > 0$,
$\sup_{x \in S} E_x|V(X_1) - V(x)| < \infty$. *Then the process cannot be positive recurrent.*

Proof. The second condition implies that $E_x|V(X_n)| < \infty$ for all $x \in S$. Now let $\tau = \inf\{n : V(X_n) < K\}$. The first condition can be written as

$$(V(X_{\tau \wedge n}), n \ge 0) \text{ is a submartingale under } P_x.$$

Hence

$$E_x V(X_{\tau \wedge n}) \ge E_x V(X_{\tau \wedge 0}) = V(x) \ge K.$$

If the process is positive recurrent then $E_x \tau < \infty$ and, by the martingale convergence theorem, $V(X_{\tau \wedge n}) \to V(X_\tau)$ in L_1. Therefore,

$$E_x V(X_\tau) \ge K.$$

But $V(X_\tau) < K$ a.s., and so we arrived at a contradiction. □

We now pass to a more general criterion. First, a definition.

Transient set A set $B \subseteq S$ is called *transient* if $P_x(\tau_B = \infty) > 0$ for all $x \in S$, where $\tau_B = \inf\{n \ge 1 : X_n \in B\}$ is the first return time to B.

By instability, here, we mean that the members of a certain class of sets is transient.

Let $L : S \to \mathbb{R}_+$ be a "norm-like" function, i.e., suppose (at least) that L is unbounded. We say that the chain is transient if each set of the form $B_N = \{x \in S : L(x) \le N\}$ is transient.

In the sequel, we will present criteria that decide whether $\lim_{n\to\infty} L(X_n) = \infty$, P_x-a.s. Clearly then, this will imply transience of each B_N.

Thinking of L as a Lyapunov function, it is natural to seek criteria that are, in a sense, opposite to those of Theorem 2. One would expect that if the drift $E_x[L(X_1) - L(X_0)]$ is bounded from below by a positive constant, outside a set of the form B_N, then that would imply instability. However, this is not true and this has been a source of difficulty in formulating a general enough criterion thus far. To the best of our knowledge, the most general criterion is Theorem 2.2.7. of Fayolle et al. (1995) which is, however, rather restrictive because (i) it is formulated for countable state Markov chains and (ii) it requires that a transition from a state x to a state y, with $L(x)-L(y)$ larger than a certain constant, is not possible. However, it gives insight as to what problems one might encounter: one needs to regulate, not only the drift from below, but also its size when the drift is large.

The theorem below is a generalization of the one mentioned above. First, define

$$\sigma_N := \tau_{B_N^c} = \inf\{n \geq 1 : \ L(X_n) > N\}$$
$$\Delta := L(X_1) - L(X_0).$$

We then have:

Theorem 4. *Suppose there exist $N, M, \varepsilon > 0$ and a measurable $h : [0,\infty) \to [1,\infty)$ with the property that $h(t)/t$ be concave-increasing on $1 \leq t < \infty$, and $\int_1^\infty h(t)^{-1} dt < \infty$, such that*

- **(I1)** $P_x(\sigma_N < \infty) = 1$ *for all* x.
- **(I2)** $\inf_{x \in B_N^c} E_x[\Delta, \Delta \leq M] \geq \varepsilon$.
- **(I3)** *The family* $\{P_x(h(\Delta) \in \cdot), \ x \in B_N^c\}$ *is uniformly integrable, i.e.,* $\lim_{K\to\infty} \sup_{x \in B_N^c} \int_K^\infty t P(h(\Delta) \in dt) = 0$.

Then $P_x(\lim_{n\to\infty} L(X_n) = \infty) = 1$, *for all* $x \in S$.

This theorem is proved in detail by Foss and Denisov (2001). We remark that there are extensions for non-homogeneous Markov chains. Condition (I1) says that the set B_N^c is recurrent. Of course, if the chain itself forms one communicating class, then this condition is automatic. Condition (I2) is the positive drift condition. Condition (I3) is the condition that regulates the size of the drift. We also note that an analog of this theorem, with state-dependent drift can also be derived. (The theorem of Fayolle et al. does use state-dependent drift.)

To see that (I3) is essential, consider the following example: Let $S := \mathbb{Z}_+$, and $\{X_n\}$ a Markov chain with transition probabilities

$$p_{i,i+1} = 1 - p_{i,0}, \quad i \geq 1,$$
$$p_{0,1} = p_{0,0} = 1/2.$$

Suppose that $0 < p_{i,0} < 1$ for all i, and $\sum_i p_{i,0} < \infty$. Then the chain forms a single communicating class. Also, with τ_0 the first return to 0, we have

$$P_i(\tau_0 = \infty) = \prod_{j \geq i} (1 - p_{j,0}) > 0.$$

So the chain is transient. However not that the natural choice for L, namely $L(x) \equiv x$ trivially makes (I2) true.

5 Topics for discussion and problems

1. Give a geometric interpretation of the Lyapunov condition $\langle V'(x), f(x)\rangle < 0$ in terms of the two vector fields $V'(x)$ and $f(x)$. Recall that $V'(x)$ is normal to a level surface of the function V.

2. Finding an appropriate Lyapunov function is an art. Discuss how the geometric picture may help you guess the form of a Lyapunov function.

3. What is the geometric picture for the van der Pol oscillator?

4. Find an example of a discrete-time deterministic recursion whose stability can be deduced by an appropriate Lyapunov function.

5. Give an example of a stable but not asymptotically stable dynamical system (e.g. an ODE).

6. Consider (ξ_n) to be i.i.d. Gaussian, say, random vectors in $\mathbb{R}^d$ with zero mean, and define the Markov chain
$$X_{n+1} = AX_n + \xi_n,$$
where A is a $d \times d$ matrix with eigenvalues having magnitude strictly smaller than 1. Show that the unit ball is positive recurrent by means of an appropriate Lyapunov function.

7. Now let $d = 1$ and, instead of A, consider a time-dependent A_n, where (A_n) are i.i.d. positive random variables:
$$X_{n+1} = A_n X_n + \xi_n.$$
Show that the unit ball is positive recurrent if $E \log A_1 < 0$.

8. Consider a Markov chain in $\mathbb{Z}_+$ with $E(X_{n+1} - X_n \mid X_n = x) \sim -c/x$, and $E((X_{n+1} - X_n)^2 \mid X_n) \to b$, as $x \to \infty$, where c, b are positive constants. Use an appropriate Lyapunov function in order to deduce that the chain is positive recurrent if $2c > b$. (Also prove that it is transient if $2c < b$. The "critical case" $2c = b$ is a tough one and what happens there depends on other conditions as well.)

9. The classical Lindley recursion is
$$X_{n+1} = (X_n + \xi_n)^+.$$
Assume that the (ξ_n) are i.i.d. integrable random variables with $E\xi_1 < 0$. Show that the set $[0, 1]$ is positive recurrent.

10. Suppose X_t is a Markov Jump Process (i.e. a Markov chain in continuous time with at most finitely many jumps on each bounded time interval, almost surely). Recall that such a process is defined (in distribution) through its transition rates $q(x, y)$, $x \neq y$. Formulate a Lyapunov function criterion directly in terms of the rates.

11. A 2-station Jackson network is a Markov chain in $\mathbb{Z}_+^2$ with $q(x, x + e_1) = \lambda$, $q(x + e_1, x + e_2) = \mu_1$, $q(x + e_2, x) = \mu_2$, $x \in \mathbb{Z}_+^2$. Here $e_1 = (1, 0)$, $e_2 = (0, 1)$ are the standard unit vectors. Find Lyapunov function when $\lambda < \mu_1 < \mu_2$ that shows that the unit ball is positive recurrent. Repeat when $\lambda < \mu_2 < \mu_1$. Hint: You may choose linear or piecewise linear functions. To do so, it is helpful to consider the geometric point of view.

12. Lyapunov functions can also be defined for continuous-time Markov processes in general state space (e.g. diffusions). Consider i.i.d. random variables $\tau_n, n \geq 1$, and let $T_n = \tau_1 + \cdots + \tau_n$. Then $X_t = t - \max\{T_n : T_n \leq t, n \geq 0\}$, $t \geq 0$, is a Markov process in $S = [0, \infty)$ known as the "age of the renewal process". The generator $\mathcal{A}$ of the process is defined as the set of pairs of measurable real functions (h, g) for which $h(X_t) - \int_0^t g(X_s)ds$ is a local martingale with respect to the filtration of the process itself. Assume $E\tau_1 = 1/\lambda \in (0, \infty)$ and that τ_1 is absolutely continuous with density f, distribution function F, and let $r = f/(1-F)$. Taking h to be a C^1 function there is essentially a unique $g := \mathcal{A}h$ for which $h(X_t) - \int_0^t g(X_s)ds$ is a local martingale. Show that $\mathcal{A}h(x) = h'(x) - (h(x) - h(0))r(x)$. Now let

$$m(x) := E(\tau_1 - x \mid \tau_1 > x), \quad V(x) := F(x)(1 + m(x)).$$

Show that, outside a compact set, $\mathcal{A}V(x) \leq -G(x)$, for some function G such that $\inf_x G(x) > 0$. This can be recognised as a Lyapunov criterion. Let $N(B) := \sum_n \mathbf{1}(T_n \in B)$. Using the results of Tuominen and Tweedie (1994), show that $\lim_{t \to \infty} E_x N(B + t) = \lambda |B|$, uniformly over all Borel ssubsets of a compact interval $[0, T]$. (See Konstantopoulos and Last (1999) for rates of convergence to this renewal theorem.)

References

Fayolle, G., Malyshev, V. and Menshikov, M. (1995). *Topics in the Constructive Theory of Markov Chains.*

Foss, S.G. and Denisov, D.E. (2001). On transience conditions for Markov chains. *Sib. Math. J.* **42** 425-433.

Foster, F.G. (1953). On the stochastic matrices associated with certain queueing processes. *Ann. Math. Stat.* **24**, 355-360.

Hirsch, M. and Smale, S. (1974). *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press.

Konstantopoulos, T. and Last, G. (1999). On the use of Lyapunov function methods in renewal theory. *Stoch. Proc. Appl.* **79**, 165-178.

Meyn, S. and Tweedie, R. (1993). *Markov Chains and Stochastic Stability.* Springer, New York.

Pakes, A.G. (1969). Some conditions for ergodicity and recurrence of Markov chains. *Oper. Res.* **17**, 1048-1061.

Tuominen, P. and Tweedie, R.L. (1994). Subgeometric rates of convergence of f-ergodic Markov chains. *Adv. Appl. Prob.* **26**.

Tweedie, R.L. (1976). Criteria for classifying general Markov chains. *Adv. Appl. Prob.* **8**, 737-771.