1 Lyapunov functions and performance approximation

In this lecture we consider the application of Lyapunov functions to obtain performance bounds. These lectures notes summarize the technical conclusions — they are not intended to be a complete record of the Lecture #3 on Lyapunov functions.

Stochastic Lyapunov theory is couched in terms of the *generator* $\mathcal{D} = P - I$. That is, for any function V, the function $\mathcal{D}V$ is defined as the mean increment,

$$\mathcal{D}V\,(x) := \mathsf{E}[V(X(t+1)) - V(X(t)) \mid X(t) = x], \qquad x \in \mathsf{X}. \tag{1}$$

The notation is used so that we can more vividly display analogies between deterministic and stochastic stability theory.

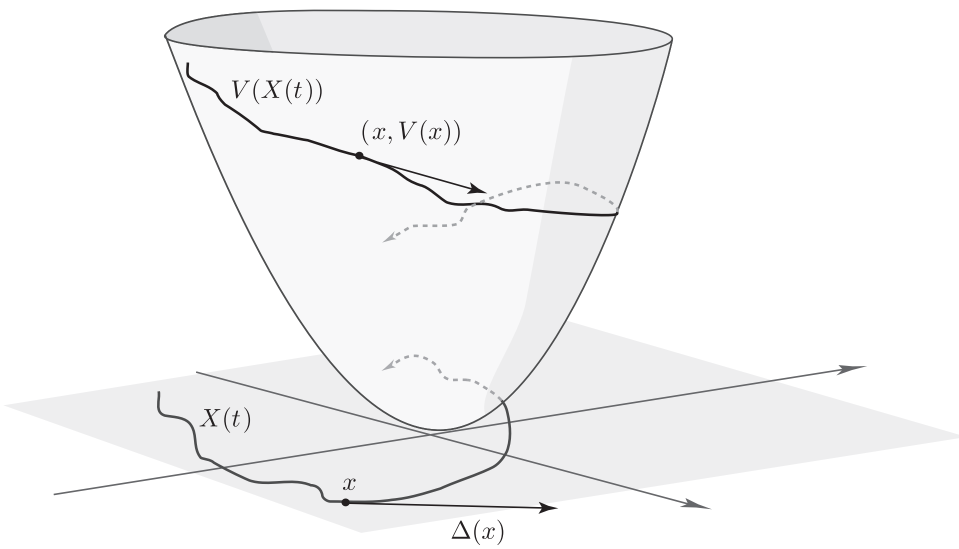

Figure 1: $V(X(t))$ is decreasing, on average, outside of the set S.

A basic stochastic Lyapunov bound is *Poisson's inequality*: for a function $V\colon \mathsf{X} \to \mathbb{R}_+$, a function $c\colon \mathsf{X} \to \mathbb{R}_+$, and a constant $\overline{\eta} < \infty$,

$$\mathcal{D}V\,(x) \le -c(x) + \overline{\eta}, \qquad x \in \mathsf{X}\,. \tag{2}$$

The function c is usually interpreted as a cost function on the state space.

Typically, it is assumed that $c(x)$ is large for "large" x, such as a norm. In this case the Poisson inequality implies that $V(X(t))$ decreases on average whenever $X(t)$ is large. This is illustrated in Figure 1, in which the set referred to in the caption is $S = \{x : c(x) \le \overline{\eta}\}$.

1.1 Lyapunov stability theory

First let's review some key aspects of Lyapunov's theory of nonlinear differential equations. Consider the dynamical system on $\mathsf{X} = \mathbb{R}^2$ defined by the nonlinear state space model,

$$\tfrac{d}{dt}x(t) = f(x(t)), \qquad x(0) = x \in \mathsf{X}. \tag{3}$$

The function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is assumed to be continuous. It is also assumed that there is a unique equilibrium x^*. Hence $f(x^*) = \mathsf{0}$, and $f(x) \neq \mathsf{0}$ for $x \neq \mathsf{0}$.

A function $V\colon \mathsf{X} \to \mathbb{R}$ is a Lyapunov function if the following conditions hold:

(i) V is non-negative valued, differentiable, and its derivative ∇V is continuous.

(ii) It is coercive: $\lim_{\|x\|\to\infty} V(x) = \infty$.

(iii) For any solution $\boldsymbol{x}$, whenever $x(t) \neq x^*$,

$$\tfrac{d}{dt} V(x(t)) < 0.$$

Naturally, $\frac{d}{dt} V(x(t)) = 0$ if $x(t) = x^*$ since in this case $V(x(t+s)) = V(x^*)$ for all $s \geq 0$.

If a Lyapunov function exists, then x^* is asymptotically stable: Hence, in particular, $x(t) \to x^*$ as $t \to \infty$ for each initial condition.[1]

This conclusion can be refined to obtain performance bounds. The drift condition (iii) can be expressed in functional form,

$$\langle \nabla V(x), f(x) \rangle < 0, \qquad x \neq x^*.$$

Suppose that $c\colon \mathsf{X} \to \mathbb{R}_+$ is a cost function that measures deviation from x^*. Assume that $c(x^*) = 0$, and that $c(x) > 0$ for $x \neq x^*$. Suppose moreover that V and c satisfy,

$$\langle \nabla V(x), f(x) \rangle \leq -c(x), \qquad x \in \mathsf{X}.$$

Returning to the differential representation, this implies that for each initial condition

$$\tfrac{d}{dt} V(x(t)) \leq -c(x(t)), \qquad t \geq 0.$$

For a given time-horizon $T > 0$, we obtain by the fundamental theorem of calculus,

$$\begin{aligned} -V(x(0)) \leq V(x(T)) - V(x(0)) &= \int_0^T \Bigl(\tfrac{d}{dt} V(x(t))\Bigr)\, dt \\ &\leq -\int_0^T c(x(t)), \qquad\qquad T \geq 0. \end{aligned}$$

On letting $T \to \infty$ we obtain a bound on the total cost,

$$\int_0^\infty c(x(t)) \leq V(x), \qquad x(0) = x \in \mathsf{X}.$$

In this part of the course we generalize these concepts to Markovian models. We will also directly apply the deterministic theory to obtain insight on the structure of solutions to Poisson's equation or inequality. One approach is to consider a *fluid model* associated with the Markov chain. In one formulation, this is expressed as the O.D.E. (3) in which f is chosen to be the *drift vector field* defined by

$$\Delta(x) := \mathsf{E}[X(t+1) - X(t) \mid X(t) = x], \qquad x \in \mathsf{X}. \tag{4}$$

[1] Asymptotic stability also requires that this convergence be uniform over the starting point $x(0)$ in any bounded subset of X.

1.2 Average cost

Let $\eta(x)$ denote the mean average cost,

$$\eta(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \mathsf{E}[c(X(t)) \mid X(0) = x] \,,$$

assuming the limit exists. The average cost is independent of x in most cases to be considered in the course. When c is a cost function, then the average cost is a natural performance metric for evaluation of the Markov chain.

The conditional expectation can be expressed in operator-theoretic notation as $P^t c\,(x)$, so that

$$\eta(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} P^t c\,(x) \,.$$

This admits a simple bound under Poisson's inequality (2).

The bound (2) can be written,

$$PV \le V - c + \overline{\eta} \tag{5}$$

Applying P to both sides then gives $P^2V \le PV - Pc + P\overline{\eta}$, and since $\overline{\eta}$ is constant,

$$P^2V \le PV - Pc + \overline{\eta} \le V - c - Pc + 2\overline{\eta}$$

By repeated multiplication by P we conclude that, for any n,

$$P^nV \le V + n\overline{\eta} - \sum_{t=0}^{n-1} P^t c$$

On rearranging terms, and using the assumption that $V \ge 0$, we obtain the following:

Proposition 1.1. *Suppose that* (2) *holds with $V \ge 0$ everywhere. Then, the following transient bound holds for each $n \ge 1$, and each $x \in \mathsf{X}$:*

$$\frac{1}{n} \sum_{t=0}^{n-1} P^t c\,(x) \le \overline{\eta} + \frac{1}{n} V(x)$$

Consequently, the average-cost bound also holds, $\eta(x) \le \overline{\eta}$.

We will see that, under very general conditions, the average-cost bound given in Proposition 1.1 is tight: There is a function $h\colon \mathsf{X} \to \mathbb{R}_+$ satisfying $\mathcal{D}h\,(x) = -c(x) + \eta(x)$ for each x.

Example: The scalar linear state space model Consider the scalar model,

$$X(t+1) = \alpha X(t) + \mathcal{E}(t+1), \qquad t \ge 0, \tag{6}$$

where $\boldsymbol{\mathcal{E}}$ i.i.d., with zero mean and finite second moment σ_e^2. The cost function is the quadratic, $c(x) = \frac{1}{2}x^2$.

Let $V(x) = \frac{1}{2}Dx^2$, with $D > 0$. We then have,

$$\begin{aligned}\mathcal{D}V\,(x) &= \mathsf{E}[V(X(t+1)) - V(X(t)) \mid X(t) = x] \\ &= \tfrac{1}{2}D\mathsf{E}[(\alpha x + \mathcal{E}(1))^2 - x^2] \\ &= \tfrac{1}{2}D(\alpha^2 - 1)x^2 + \tfrac{1}{2}D\sigma_e^2\end{aligned} \tag{7}$$

Provided $|\alpha| < 1$, we can set $D = (1-\alpha^2)^{-1}$ in the definition of V to obtain a solution to Poisson's *equation* with forcing function c,

$$\mathcal{D}V\,(x) = -c(x) + \overline{\eta} \tag{8}$$

with $\overline{\eta} = \frac{1}{2}(1-\alpha^2)^{-1}\sigma_e^2$.

It follows from Proposition 1.1 that $\eta(x) \le \overline{\eta}$ for each x. In fact, the steps above show that $\mathsf{E}[c(X(t))] \to \overline{\eta}$ for each initial condition, so that we have equality: $\eta(x) \equiv \overline{\eta}$.

Note that if $\boldsymbol{\mathcal{E}}$ is Gaussian then $x(t)$ has a Gaussian distribution for each t. In this case it is clear that $P^t(x,\,\cdot\,)$ converges to a Gaussian $N(0,\sigma_\infty^2)$ distribution as $t \to \infty$, with $\sigma_\infty^2 = (1-\alpha^2)^{-1}\sigma_e^2$.

1.3 Discounted cost

In economics and operations research applications the discounted cost criterion is typically favored. Given a discount parameter $\beta \in (0,1)$, the discounted cost from initial condition x is defined as the weighted sum,

$$h_\beta(x) = \sum_{t=0}^{\infty} \beta^t \mathsf{E}[c(X(t)) \mid X(0) = x]\,.$$

Once again this has the operator-theoretic form,

$$h_\beta = \sum_{t=0}^{\infty} \beta^t P^t c\,. \tag{9}$$

If c is non-negative valued then the lower bound $h_\beta(x) \ge c(x)$ holds, so that the discounted cost is unbounded whenever this is true of c. And, once again, we obtain a bound on h_β under Poisson's inequality.

The bound (2), expressed in the form (5), gives

$$\beta P V \le V - g + \beta\overline{\eta} \tag{10}$$

where $g = (1-\beta)V + \beta c$, a convex combination of the Lyapunov function and the cost function. Applying βP to each side gives,

$$(\beta P)^2 V \le \beta P V - \beta P g + \beta^2\overline{\eta}$$

and then using (10),

$$(\beta P)^2 V \le V - g - \beta P g + \beta\overline{\eta} + \beta^2\overline{\eta}$$

As in the average-cost problem we obtain by induction,

$$(\beta P)^n V \le V - \sum_{t=0}^{n-1} \beta^t P^t g + \overline{\eta} \sum_{t=0}^{n-1} \beta^{t+1}$$

From the definition $g = (1-\beta)V + \beta c$ we obtain,

Proposition 1.2. *If (2) holds with $V \ge 0$ everywhere then for each $\beta \in (0,1)$ and for each $x \in \mathsf{X}$:*

$$\sum_{t=0}^{\infty} \beta^t \big(\beta P^t c\,(x) + (1-\beta) P^t V\,(x)\big) \le V(x) + \frac{\beta}{1-\beta}\overline{\eta}. \tag{11}$$

Consequently, the discounted cost satisfies $h_\beta(x) \le V(x) + \overline{\eta}(1-\beta)^{-1}$ for each x.

The final bound deserves some additional explanation. Using the fact that $V \ge 0$ we can drop all but one term involving V on the left hand side of (11) to obtain,

$$(1-\beta)V(x) + \beta \sum_{t=0}^{\infty} \beta^t P^t c\,(x) \le V(x) + \frac{\beta}{1-\beta}\overline{\eta}.$$

The bound on $h_\beta(x)$ follows on adding $(1-\beta)V(x)$ to each side, and then dividing each side of the resulting bound by β.

The bound obtained in Proposition 1.2 is not tight in general. To see why, note that the function $h_\beta(x)$ solves the dynamic programming equation,

$$h_\beta(x) = c(x) + \beta \mathsf{E}[h_\beta(X(1)) \mid X(0) = x] \tag{12}$$

or, in operator notation, $h_\beta = c + \beta P h_\beta$. The proof of Proposition 1.2 can be interpreted as an approximation to this equation based on a scaled multiple of V.

Example: Discounted cost for the scalar linear state space model Poisson's equation (8) can be written,

$$PV\,(x) = V(x) - c(x) + \overline{\eta}\,.$$

Proposition 1.1 gives the bound $h_\beta \le V + \overline{\eta}(1-\beta)^{-1}$, or

$$h_\beta(x) \le \tfrac{1}{2}(1-\alpha^2)^{-1}\Big(x^2 + (1-\beta)^{-1}\sigma_e^2\Big) \tag{13}$$

To compute h_β we apply the dynamic programming equation (12). Let $V(x) = A_\beta + \frac{1}{2}D_\beta x^2$, with A_β, D_β constants to be chosen. From (7) we have,

$$\mathcal{D}V = -(1-\alpha^2)(V - A_\beta) + \tfrac{1}{2}D_\beta \sigma_e^2$$

or $PV \le \alpha^2 V + (1-\alpha^2)A_\beta + \frac{1}{2}D_\beta\sigma_e^2$. Scaling by β and adding c to each side gives,

$$c + \beta PV = c + \beta\big(\alpha^2 V + (1-\alpha^2)A_\beta + \tfrac{1}{2}D_\beta\big)$$

Or, re-introducing the quadratic expressions,

$$c(x) + \beta PV(x) = \tfrac{1}{2}(1 + \beta\alpha^2 D_\beta)x^2 + \beta(\tfrac{1}{2}D_\beta\sigma_e^2 + A_\beta)$$

To solve the dynamic programming equation we require that the right hand side coincide with V. This requires the solution of two equations:

$$1 + \beta\alpha^2 D_\beta = D_\beta \qquad \beta(\tfrac{1}{2}D_\beta\sigma_e^2 + A_\beta) = A_\beta$$

Solving for D_β and A_β gives,

$$h_\beta(x) = A_\beta + \tfrac{1}{2}D_\beta x^2 = \tfrac{1}{2}(1 - \beta\alpha^2)^{-1}\big(x^2 + \beta\sigma_e^2\big)$$

In particular, the bound (13) does indeed hold.

Next steps: Finer conclusions based on a Lyapunov function.