1 Why do I need probability in a deterministic world?

Deterministic systems can be unpredictable in the classical sense and require a probabilistic description. In this section we elaborate on this observation.

In 1963, E. N. Lorenz introduced the following system of ODEs:

$$\begin{cases} \dot{x} = \sigma(y - x), \\ \dot{y} = x(r - z) - y, \\ \dot{z} = xy - bz, \end{cases} \tag{1}$$

where σ, r, b are positive parameters. (1) has become the paradigm of a system displaying *chaos*, defined as *aperiodic long-time behavior in a deterministic system which exhibits sensitive dependence on initial conditions*. This definition contains three main elements:

1. *Aperiodic long-time behavior,* meaning that generic bounded trajectories do not settle on fixed points or periodic orbits.

2. *Deterministic,* meaning that there is no random input in the system.

3. *Sensitive dependence on initial conditions,* meaning that nearby trajectories separate exponentially fast in time, i.e. the system has a positive Lyapunov exponent.

The standard parameter setting for studying Lorenz system in the chaotic regime is $\sigma = 10$, $r = 28$, $b = 8/3$. Numerical experiments conducted at this values confirm the properties above for the solutions of (1). All trajectories of (1) eventually settle a set of zero volume in phase-space, referred to as the Lorenz attractor and shown on figure 1. Lorenz attractor has a characteristic butterfly shape and the trajectories spiral outward on one wing before making a transition to the vicinity of the center of the other wing, and repeating the process – see figure 2. The sensitive dependence to initial conditions is apparent here from the fact that two nearby trajectories eventually jump at different times from one wing to the other, as illustrated on figure 2.

In the original paper, Lorenz noted the following remarkable fact. Let m_n be the value of the nth maximum of z. Plotting in the plane the successive pairs (m_n, m_{n+1}) for $n = 1, 2, \ldots$ produces a figure where the points (m_n, m_{n+1}) line up on a graph with a tent shape. Therefore it is as if the solutions of (1) lead to a symbolic dynamics for m_n,

$$m_{n+1} = f(m_n),$$

with the function f whose graph is shown in figure 3. Generally, given a function f, a recurrence relation like this one defines what is called a one-dimensional map. These are much simpler to analyze than ODEs like (1), and we now use

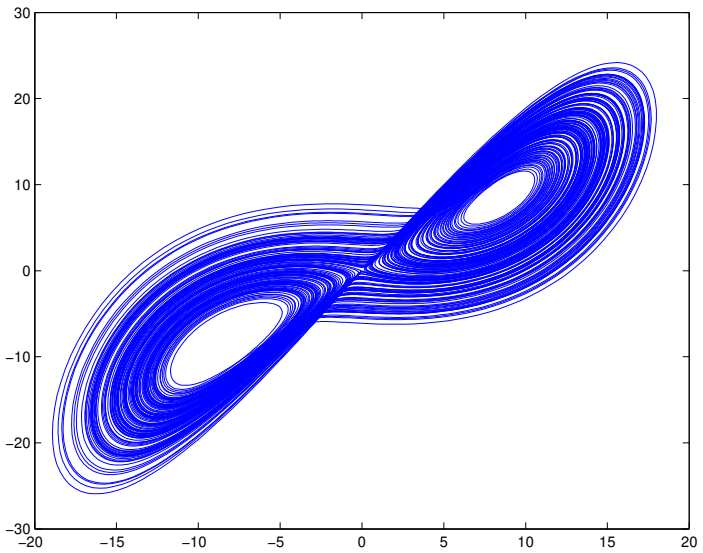

Figure 1: A given trajectory in the (x, y) plane. The trajectory settles on the Lorenz attractor.

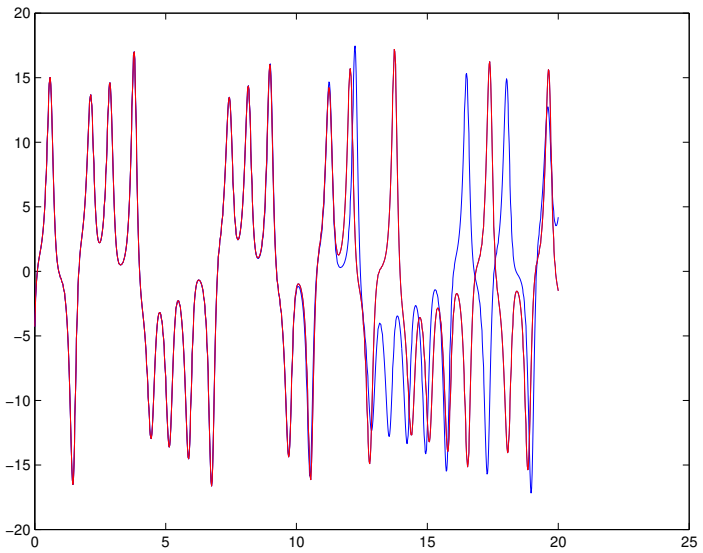

Figure 2: The x-component of two trajectories whose initial conditions differ by 10^{-4}.

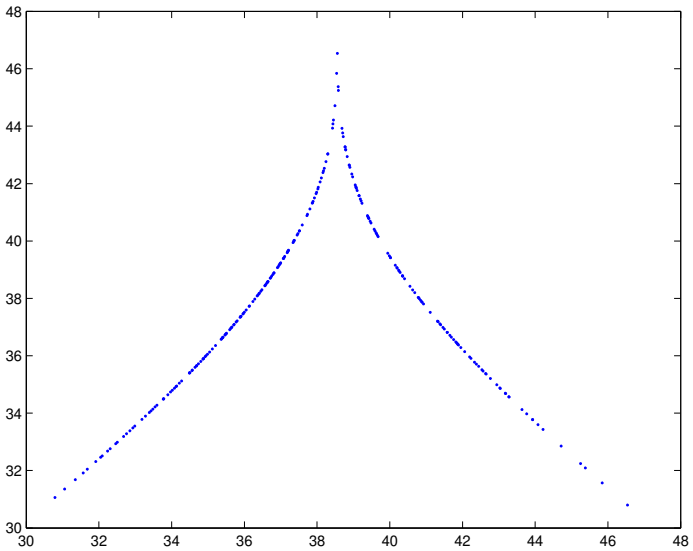

Figure 3: The successive pairs (m_n, m_{n+1}) for $n = 1, 2, \ldots$. Thus the dynamics of m_n seems reducible to that of a one-dimensional map.

a map similar (up to rescaling) to the one for m_n where the function f is given by

$$f(x) = \begin{cases} 2x & \text{if } \ x \leq 1/2 \\ 2 - 2x & \text{if } \ x > 1/2. \end{cases}$$

If $x_0 \in [0, 1]$, the recurrence relation

$$x_{n+1} = f(x_n), \tag{2}$$

then defines a dynamics on the interval $[0, 1]$ – see figure 4. This map is referred to as the tent map and it is not difficult to analyze its properties – in view of the analogy with the dynamics for m_n, these will help elucidate some of the properties of Lorenz system.

Recall that to every number $x \in [0, 1]$ one can associated a sequence $\{\alpha_j\}_{j\in\mathbb{N}}$, with $\alpha_j = 0$ or $\alpha_j = 1$, such that

$$x = \sum_{j\in\mathbb{N}} \alpha_j 2^{-j}.$$

The sequence $\{\alpha_j\}_{j\in\mathbb{N}}$ is called the binary representation of x, and we shall write $x = [\alpha_1, \alpha_2, \ldots]$. Now suppose the initial condition x_0 for the map in (2) is

$$x_0 = [\alpha_1, \alpha_2, \ldots].$$

If $\alpha_1 = 0$, then $x_0 < 1/2$, and it follows that

$$x_1 = [\alpha_2, \alpha_3, \ldots].$$

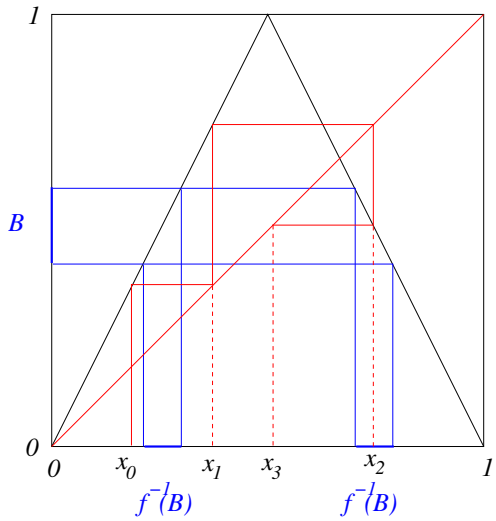

Figure 4: The tent map. The figure shows the first three iterates of x_0, and the pre-image of the set B.

Similarly if $\alpha_1 = 1$, then $x_0 \geq 1/2$, and it follows that

$$x_1 = [N\alpha_2, N\alpha_3, \ldots],$$

where N is the negation operator, defined so that $N0 = 1$ and $N1 = 0$. If one agrees that $N^0 = I$ (identity operator), these two expression can be combined into

$$x_1 = [N^{\alpha_1}\alpha_2, N^{\alpha_1}\alpha_3, \ldots].$$

Iterating the argument gives

$$x_n = [N^{\alpha_1+\alpha_2+\cdots+\alpha_n}\alpha_{n+1}, N^{\alpha_1+\alpha_2+\cdots+\alpha_n}\alpha_{n+2}, \ldots], \tag{3}$$

which gives an explicit expression for the nth iterate of the map.

(3) shows that most trajectories in the tent map are aperiodic. Indeed, only the rational numbers in $[0, 1]$, i.e. these numbers for which

$$x = [\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \beta_1, \ldots, \beta_q, \ldots]$$

for some $p, q \in \mathbb{N}$, are such that they eventually settle to the only fixed point $x = 0$, or to a periodic orbit (both unstable), and these numbers form a zero measure set. All the irrational numbers in $[0, 1]$, one the other hand, lead to aperiodic orbits. (3) also shows that the solution of the tent map exhibits sensitive dependence on initial conditions. Consider indeed two initial conditions, x_0 and y_0 such that

$$|x_0 - y_0| \leq 2^{-N}.$$

for some arbitrary $N \in \mathbb{N}$. Thus the first N binary digits of x_0 and y_0 are identical, and they may differ only afterwards, i.e. they are very close if N is large. Yet, it is immediate from (3) that the bound above deteriorates rapidly in time and, in particular, one can only conclude that

$$|x_n - y_n| \leq 1$$

as soon as $n > N$.

This indicates that, despite the deterministic nature of the tent map, it is illusory to try to make long-term predictions about the solutions because small error in the initial conditions are inevitable but eventually dominate the solution. Errors in initial conditions can be represented probabilistically. Therefore, a probabilistic description of the tent map seems appropriate, and we focus on this next.

Suppose that the initial condition of the tent map is a random variable with distribution μ_0, i.e. for every set $B \in [0, 1]$,

$$\mu_0(B) = \text{Prob}\{x_0 \in B\},$$

and of course $\mu([0, 1]) = 1$. This may be a way to represent errors in initial conditions. For instance if our measurement of the initial condition gives x, but there is a possible error δ on this measurement, one may take the uniform distribution on $[x - \delta, x + \delta]$ for μ_0.

How does μ_0 evolve by the tent map? In other words, if μ_0 is the probability distribution of x_0, and $x_1 = f(x_0)$, what is the distribution μ_1 of x_1? The answer is easily found from the definition:

$$\mu_1(B) = \text{Prob}\{x_1 \in B\} = \text{Prob}\{f(x_0) \in B\}.$$

Denote by $f^{-1}(B)$ the pre-image of B, i.e. the set such that $f(f^{-1}(B)) = B$ – see figure 4. Then

$$\mu_1(B) = \text{Prob}\{x_0 \in f^{-1}(B)\} = \mu_0(f^{-1}(B)).$$

It is straightforward to iterate the argument. If μ_n is the the probability distribution of x_n, the nth iterate of the map, then

$$\mu_n(B) = \mu_0(f^{-n}(B)).$$

Now, what do we gain from this viewpoint? Well, while the evolution of x_n is aperiodic and unpredictable, the evolution of μ_n is simpler. In fact, it is easy to show (do it!) that if μ_0 is continuous, then μ_n converges to the uniform distribution on $[0, 1]$ as $n \to \infty$, i.e.

$$\mu_n \to \mu \quad \text{as} \quad n \to \infty \quad \text{with } \mu(B) = \text{length}(B). \tag{4}$$

(In particular, $\mu(dx) = dx$.) The limiting distribution μ in (4) is called the *invariant measure* of the map. It is the unique continuous measure which is left invariant by the map, i.e.

$$\mu(B) = \mu(f^{-1}(B)).$$

The importance of the invariant measure is made apparent from *Birkhoff ergodic theorem*, which states that, given any test function η on $[0,1]$, one has

$$\frac{1}{n}\sum_{j=0}^{n-1}\eta(x_j) \to \int_0^1 \eta(x)dx \quad \text{as} \quad n \to \infty,$$

for almost all initial conditions x_0 (i.e. except for a set of Lebesgue measure zero).