Introduction to Lyapunov Spectrum

The Lyapunov spectrum is a set of numbers used to characterize the stability and chaotic behavior of dynamical systems. Named after the Russian mathematician Aleksandr Lyapunov, it provides insights into how trajectories diverge or converge in phase space over time.

Consider a dynamical system described by a map or flow $f : \mathbb{R}^d \to \mathbb{R}^d$. Let x_0 be a point in the phase space of the system. The Lyapunov spectrum consists of a set of Lyapunov exponents, denoted by λ_i, which quantify the rate of exponential growth or decay of infinitesimally close trajectories.

For each i from 1 to d, the i-th Lyapunov exponent λ_i is defined as the limit:

$$\lambda_i = \lim_{n\to\infty} \frac{1}{n} \log |\det(Df^n(x_0))| \tag{1}$$

where $Df^n(x_0)$ represents the Jacobian matrix of f evaluated at x_0 after n iterations.

The Lyapunov exponents measure the average rate of expansion or contraction of nearby trajectories along different directions in the phase space. A positive Lyapunov exponent $\lambda_i > 0$ indicates divergence, suggesting chaotic behavior, while a negative exponent $\lambda_i < 0$ suggests convergence towards a stable attractor.

The Lyapunov spectrum provides a comprehensive description of the system's behavior by characterizing the stability along all possible directions in phase space. By analyzing these exponents, researchers can understand the underlying dynamics of complex systems and predict their long-term behavior.

Example of Lyapunov Spectrum

The Lyapunov spectrum is a set of numbers used to characterize the stability and chaotic behavior of dynamical systems. Named after the Russian mathematician Aleksandr Lyapunov, it provides insights into how trajectories diverge or converge in phase space over time.

Consider a simple 2-dimensional dynamical system defined by the map:

$$f(x, y) = (2x - y^2, x + y) \tag{2}$$

Let's choose a point in the phase space, $x_0 = (1, 1)$, and calculate the Lyapunov spectrum at this point. The Jacobian matrix of f evaluated at $(1, 1)$ is:

$$Df(1,1) = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix} \tag{3}$$

To compute the Lyapunov exponents, we iteratively apply Df to a set of initial vectors. Let's start with the initial vector $v_0 = (1, 0)$:

$$v_1 = Df(1,1) \cdot v_0 = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \tag{4}$$

The Lyapunov exponent λ_1 corresponding to the direction of v_0 is then:

$$\lambda_1 = \lim_{n\to\infty} \frac{1}{n} \log |Df^n(1,1) \cdot v_0| \tag{5}$$

$$= \lim_{n\to\infty} \frac{1}{n} \log | \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}^n \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} | \tag{6}$$

Similarly, we can choose another initial vector, say $v_0 = (0, 1)$, and calculate the Lyapunov exponent λ_2 corresponding to this direction.

By repeating this process for different initial vectors, we can obtain the complete Lyapunov spectrum for the dynamical system.

Lyapunov Spectrum in Backpropagation

In the context of backpropagation, we consider the following iterative dynamics:

$$x_{n+1} = f(x_n) \tag{7}$$

$$v_{n+1} = Df(x_n)v_n \tag{8}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ is a differentiable function and $Df(x_n)$ denotes the Jacobian matrix of f evaluated at x_n.

We define a sequence of vector spaces $V_1 \subset V_2 \subset V_3 \subset \cdots \subset T_{x_0}M$, such that for each x_0 and $v \in V_i \backslash V_{i-1}$, we have:

$$\lim_{n\to\infty} \frac{1}{n} \log |Df^n(x_0)v| = \lambda_i \tag{9}$$

where λ_i represents the i-th Lyapunov exponent. The Lyapunov spectrum provides valuable information about the stability and chaotic behavior of the dynamical system described by f. Each Lyapunov exponent λ_i corresponds to the local exponential growth rate of trajectories along the corresponding direction in the tangent space.

Calculating Lyapunov Spectra

First, we choose d random vectors:

$$[e_{1,0}, e_{2,0}, \cdots, e_{d,0}] = e_0 \tag{10}$$

where $e_{i,0} \in \mathbb{R}^d$ and $e_{i,0} \neq 0$ for all i. Then we calculate e iteratively using:

$$e_{n+1} = Df(x_n)e_n \tag{11}$$

Let T be such that $e_T = QR$, where Q is an orthogonal matrix and R is an upper triangular matrix. Then we assume:

$$Q = LV_1 \tag{12}$$

$$R = LE \tag{13}$$

where L is a lower triangular matrix, V_1 is a diagonal matrix, and E is an upper triangular matrix. We can calculate V_1 as follows:

Finally, we obtain

$$Df^{NA}e = QR \tag{14}$$

$$= Q_N R_N R_{N-1} \cdots R_1 \tag{15}$$

Then we have

$$L_1 V_1 = Q_N (R_A R_{A-1} \cdots R_1) \tag{16}$$

$$L_1 E_1 \approx \frac{1}{NA} \log \operatorname{diag}(R_A R_{A-1} \cdots R_1) \tag{17}$$

A Result

Let $\epsilon_n = Df^T(x_n)\epsilon_{n+1}$, then we have

$$\langle \epsilon_n, e_n \rangle = \langle \epsilon_{n+1}, e_{n+1} \rangle \tag{18}$$

Hence,

$$\epsilon_0 \cdot e_0 = (Df^T)^N \epsilon_N \cdot e_0 = \epsilon_N \cdot e_N = \epsilon_N \cdot (Df^N) e_0 \tag{19}$$

We only consider the first few terms of ϵ and e:

$$1 = \epsilon_0 \cdot e_0 = \epsilon \cdot Df^N(e_0) \tag{20}$$

$$= |\epsilon_{0,u} \cdot \frac{e_0}{|e_0|}| = |\epsilon_{N,u} \cdot \frac{Df^N(e_{0,u})}{|e_{0,u}|}| = |\epsilon_{N,u} \frac{|Df^N e_{0,u}|}{|e_{0,u}|}| \tag{21}$$