Lyapunov Spectrum in Backpropagation

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

$$x_{n+1} = f(x_n) \tag{1}$$
$$v_{n+1} = Df(x_n)v_n \tag{2}$$

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:

$$\lim_{n\to\infty} \frac{1}{n} \log |Df^n(x_0)v| = \lambda_i \tag{3}$$

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:

$$[e_{1,0}, e_{2,0}, \cdots, e_{d,0}] = e_0 \tag{4}$$

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

$$e_{n+1} = Df(x_n)e_n \tag{5}$$

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

$$Q = LV_1 \tag{6}$$
$$R = LE \tag{7}$$

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

$$
\begin{aligned}
Df^{NA}e &= QR && (8)\\
&= Q_N R_N R_{N-1} \cdots R_1 && (9)
\end{aligned}
$$

Then we have

$$
\begin{aligned}
L_1 V_1 &= Q_N (R_A R_{A-1} \cdots R_1) && (10)\\
L_1 E_1 &\approx \frac{1}{NA} \log \operatorname{diag}(R_A R_{A-1} \cdots R_1) && (11)
\end{aligned}
$$

A Result

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

$$
\langle \epsilon_n, e_n \rangle = \langle \epsilon_{n+1}, e_{n+1} \rangle \tag{12}
$$

Hence,

$$
\epsilon_0 \cdot e_0 = (Df^T)^N \epsilon_N \cdot e_0 = \epsilon_N \cdot e_N = \epsilon_N \cdot (Df^N) e_0 \tag{13}
$$

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

$$
\begin{aligned}
1 &= \epsilon_0 \cdot e_0 = \epsilon \cdot Df^N(e_0) && (14)\\
&= |\epsilon_{0,u} \cdot \frac{e_0}{|e_0|}| = |\epsilon_{N,u} \cdot \frac{Df^N(e_{0,u})}{|e_{0,u}|}| = |\epsilon_{N,u} \frac{|Df^N e_{0,u}|}{|e_{0,u}|}| && (15)
\end{aligned}
$$