Lecture 7

Orthogonal initialization

$$h_t = \cancel{\sigma}\left(\cancel{W_{xh}^T x_t} + W_{hh}^T h_{t-1} + \cancel{b}\right)$$

$$= W_{hh}^T h_{t-1}$$

$$= W h_{t-1}$$

$$h_n = W(\dots(W(W h_0))\dots)$$

$$= W^n h_0$$

$$< W^n$$

Fibonacci sequence

$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots$

$$\begin{pmatrix} f_{k+2} \\ f_{k+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_{k+1} \\ f_k \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

if $f_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then $f_{k+1} = A f_k$

$f_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\qquad = A^{k+1} f_0$

$$A = S \Lambda S^{-1}$$

where $S = \begin{pmatrix} \varphi & -\varphi^{-1} \\ 1 & 1 \end{pmatrix}$

$$\Lambda = \begin{pmatrix} \varphi & 0 \\ 0 & -\varphi^{-1} \end{pmatrix}$$

$$A^n = (S \Lambda S^{-1})^n$$

$$= S \Lambda^n S^{-1}$$

$$\Lambda^n = \begin{pmatrix} (\varphi)^n & 0 \\ 0 & (-\varphi^{-1})^n \end{pmatrix}$$

<u>Thm:</u>

Let $\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \dots |\lambda_d|\}$

if $\rho(A) < 1$, then $\lim_{n \to \infty} \|A^n\| = 0$

if $\rho(A) > 1$, then $\lim_{n \to \infty} \|A^n\| = +\infty$

$\Rightarrow$ Set W such that $W^n = S \Lambda^n S^{-1}$ does not vanish nor explode

$\Rightarrow$ Have $\Lambda = \begin{pmatrix} \pm 1 & & & \\ & \pm 1 & & \\ & & \pm 1 & \\ & & & \ddots \end{pmatrix}$

$\Rightarrow$ Force W to be orthogonal.