

ディープラーニングと物理学

4.2 再帰的ニューラルネットワークと誤差逆伝播法

須賀勇貴

茨城大学大学院 理工学研究科 量子線科学専攻 2 年

April 15, 2023

時系列データについて

系列データ

個々の要素が順序付きの集まりとして与えられるデータのこと
(ex)

- 動画データ → 順序付きの自然画像データ
- 文章データ → 順序付きの文字画像データ
- 会話データ → 順序付きの音声データ

長さが T の系列データは以下のように表現できる

$$\begin{pmatrix} x(1) \\ x(2) \\ x(3) \\ \vdots \\ x(T) \end{pmatrix} = |x(t)\rangle \quad (t = 1, 2, 3, \dots, T)$$

時系列データについて

(ex) "This is an apple ." という文章データを系列データとして扱う場合

$$|x(1)\rangle = |\text{This}\rangle$$

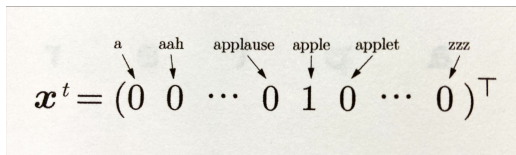
$$|x(2)\rangle = |\text{is}\rangle$$

$$|x(3)\rangle = |\text{an}\rangle$$

$$|x(4)\rangle = |\text{apple}\rangle$$

$$|x(5)\rangle = |.\rangle$$

文字データなどは 1-of-K ベクトルなどにより数値ベクトルとして表現される



The diagram illustrates the 1-of-K vector representation for the word "apple". It shows a vector x^t as a row of elements in parentheses, followed by a superscript T . The elements are: 0, 0, an ellipsis, 0, 1, 0, an ellipsis, and 0. Above each element is a label with an arrow pointing to it: 'a' points to the first 0, 'aah' points to the second 0, 'applause' points to the first 0 after the ellipsis, 'apple' points to the 1, 'applet' points to the 0 after the 1, and 'zzz' points to the final 0.

$$x^t = (0 \quad 0 \quad \cdots \quad 0 \quad 1 \quad 0 \quad \cdots \quad 0)^T$$

図: "apple" の 1-of-k ベクトル表示

再帰的ニューラルネットワーク (RNN) の考え方

これまで、扱ってきたデータはデータ間につながりがないものだった



系列データをニューラルネットワークで扱えるようにしたい



データ間のつながりを表現できるようなニューラルネットワークを構築すれば
よい！

再帰的ニューラルネットワーク (RNN) の考え方

素朴な考え方

⇒ 前の時刻の出力を次の時刻の入力に加えるようなニューラルネットワーク

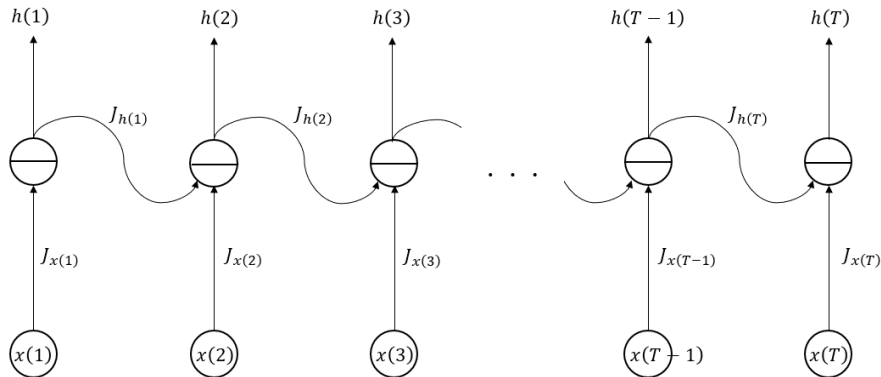


図: 最もシンプルな形の RNN

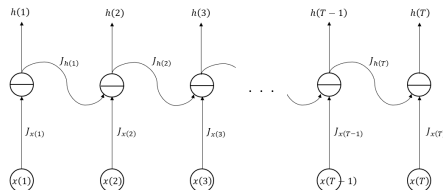
再帰的ニューラルネットワーク (RNN) の考え方

$$h(1) = \sigma_{\bullet}(J(1)x(1))$$

$$h(2) = \sigma_{\bullet}(J(2)x(2) + J_{h(1)}h(1))$$

⋮

$$h(t) = \sigma_{\bullet}(J(t)x(t) + J_{h(t-1)}h(t-1))$$



右のブラケット表記を用いると

$$\begin{aligned} |h(t)\rangle &= \sigma_{\bullet}(J(t) |x(t)\rangle + J_{h(t-1)} |h(t-1)\rangle) \\ &= \sigma_{\bullet} \sum_m |m\rangle \langle m| J(t) |x(t)\rangle \\ &\quad + \langle m| J_{h(t-1)} |h(t-1)\rangle) \end{aligned}$$

$$|x(t)\rangle = (x(1), x(2), \dots, x(T))^{\top}$$

$$|h(t)\rangle = (h(1), h(2), \dots, h(T))^{\top}$$

$$\mathbf{J}_x = \text{diag}(J_{x(1)}, J_{x(2)}, \dots, J_{x(T)})$$

$$\mathbf{J}_h = \text{diag}(J_{h(1)}, J_{h(2)}, \dots, J_{h(T)})$$

再帰的ニューラルネットワーク (RNN) の考え方

各時刻での出力は以下のように簡単に求められる

$$h(1) = \sigma_{\bullet}(J(1)x(1))$$

$$h(2) = \sigma_{\bullet}(J(2)x(2) + J_{h(1)}h(1))$$

$$\vdots$$

$$h(t) = \sigma_{\bullet}(J(t)x(t) + J_{h(t-1)}h(t-1))$$

右のブラケット表記を用いると

$$\begin{aligned} |h(t)\rangle &= \sigma_{\bullet}(J(t)|x(t)\rangle + J_{h(t-1)}|h(t-1)\rangle) \\ &= \sigma_{\bullet} \sum_m |m\rangle \left(\langle m| J(t) |x(t)\rangle + \langle m| J_{h(t-1)} |h(t-1)\rangle \right) \end{aligned}$$