

# 深層学習 day3 section3

Q) GRUもLSTMと同様にRNNの一種であり、

単純なRNNにおいて、問題となる勾配消失問題を解決し、長期的な依存関係を学習することができる。LSTMに比べて単純なモデル。

以下のPGはGRUの順伝播を行うPGである。

(こ)にあてはまるのはどれか。但し、

-sigmoid関数は要素ごとにシグモイド関数を用いる

def gru(x, h, w\_r, U\_r, W\_z, U\_z, W, U):

# ゲートを計算

r = -sigmoid(x.dot(W\_r.T) + h.dot(U\_r.T))

z = -sigmoid(x.dot(W\_z.T) + h.dot(U\_z.T))

h\_bar = np.tanh(x.dot(W.T) + (r \* h).dot(U.T))

h\_new = (=)

return h\_new

Ans) (4)  $(1-z) * h + z * h\_bar$

∵ 図上の  $z(t) \cdot h(t-1) + (1-z(t)) \cdot h(t)$  に対応