

سوال 1-

الف) اگر z_t صفر باشد، اثر حالت قبلی (state قبلی) در state جدید در نظر گرفته نمی شود (در واقع نسبتی از state قبلی است که به حالت جدید اضافه می شود). (میزان ارزش دهی به حالت قبلی است) در واقعیت update است.

r_t میزان ارزش state قبلی در تعیین state جدید را تعیین می کند. در واقع حالت جدید تابعی از ترکیب خطی ورودی و حالت قبلی است. r_t مشخص می کند که اثر حالت قبلی در مقابل ورودی چقدر است. در واقعیت reset است.

ب) طبق رابطه موجود در اسلاید 1

$$\frac{\partial L}{\partial W_h} = \sum_{j=0}^{T-1} \sum_{k=1}^j \frac{\partial y_j}{\partial y_j} \frac{\partial y_j}{\partial h_j} \left(\prod_{m=k+1}^j \frac{\partial h_m}{\partial h_{m-1}} \right) \frac{\partial h_k}{\partial W_h}$$

$$h_t = z_t \circ h_{t-1} + (1-z_t) \circ \tilde{h}_t \quad \tilde{h}_t = \tanh(W^{hx} x_t + r_t \circ W^{kh} h_{t-1})$$

$$\frac{\partial h_m}{\partial h_{m-1}} = \text{diag}(z_t) + \frac{\partial (1-z_t) \circ \tilde{h}_m}{\partial h_{m-1}} \Rightarrow \frac{\partial (1-z_t) \circ \tilde{h}_m}{\partial h_{m-1}} = \underbrace{\frac{\partial (1-z_t) \circ \tilde{h}_m}{\partial \tilde{h}_m}}_{\text{diag}(1-z_t)} \cdot \frac{\partial \tilde{h}_m}{\partial h_{m-1}}$$

$$= \text{diag}(1-z_t) \frac{\partial \tilde{h}_m}{\partial h_{m-1}}$$

$$\Rightarrow \frac{\partial h_m}{\partial h_{m-1}} = \text{diag}(z_t) + \text{diag}(1-z_t) \frac{\partial \tilde{h}_m}{\partial h_{m-1}}$$

$$\frac{\partial h_m}{\partial h_{m-1}} = I$$

اگر $z_t = 1$ باشد، در این صورت

در نتیجه ضرب های متوالی تبدیل به I می شوند، gradient vanishing، gradient exploding نخواهیم داشت.